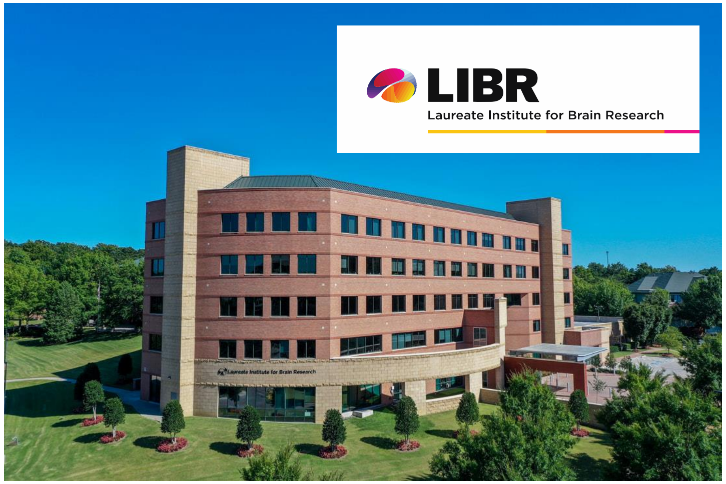
AN INTRODUCTION TO ACTIVE INFERENCE

Ryan Smith

Laureate Institute for Brain Research

Background

- Laureate Institute for Brain Research
  - Non-profit psychiatry research institute
  - Oklahoma, USA



- My lab at LIBR focuses on **using active inference** models to help **understand** the neurocognitive processes underlying **psychiatric disorders**

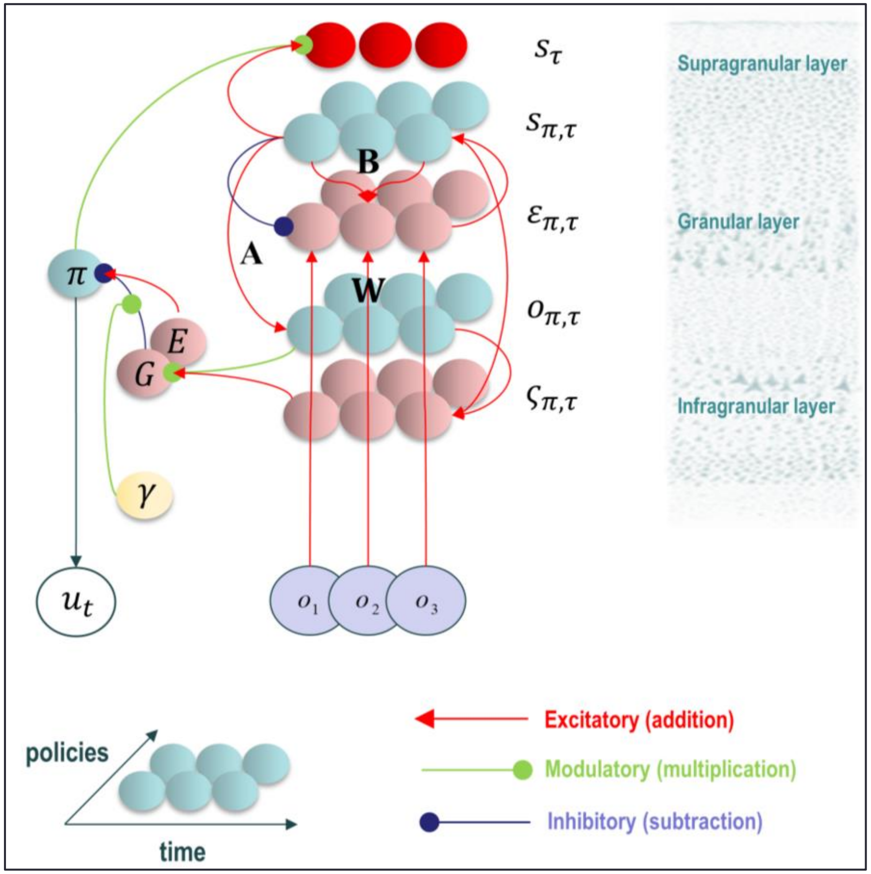
- I’m hoping to bring out why these models can be useful in **empirical research**

What is active inference good for?

- Before I tell you what active inference is, I'm going to **tell you what it's good for**.
  - It incorporates **perception**, **learning**, and **decision-making** within a **single model**
    - The framework can be applied to a **wide range of problems**
      - **Equations** used for inference are **generic** across generative model architectures
  - The **task is to figure out the right generative model structure** to simulate a particular cognitive process or behavioral task

What is it good for?

- Motivated by **biological plausibility**
  - Clear ways in which **neural networks can implement all the linked equations** in the model
- It has an accompanying **neural process theory**
  - Provides the opportunity to generate and test **precise hypotheses** in **neuroimaging studies**
    - EEG - event-related potentials (**ERPs**)
    - fMRI - **localized neural responses**



What is it good for?

- It provides a **unique approach** for modelling **explore-exploit trade-offs**
  - When do I seek reward?
  - When do I first seek out information?

- Useful framework for modelling **behavioral tasks** involving **information-seeking** and **planning**

Empirical Studies

- **On Information-seeking:**

Learning from negative outcomes in substance use disorders

ELSEVIER

Drug and Alcohol Dependence
Volume 215, 1 October 2020, 108208

Imprecise action selection in substance use disorder: Evidence for active learning impairments when solving the explore-exploit dilemma

Ryan Smith [a], Philipp Schwartenbeck [b], Jennifer L. Stewart [a], Rayus Kuplicki [a], Hamed Ekhtiari [a], Martin P. Paulus [a], Tulsa 1000 Investigators [1]

Accounting for patterns in selective attention

PLOS ONE

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

Human visual exploration reduces uncertainty about the sensed world

M. Berk Mirza, Rick A. Adams, Christoph Mathys, Karl J. Friston

Empirical Studies

- **On Planning**:

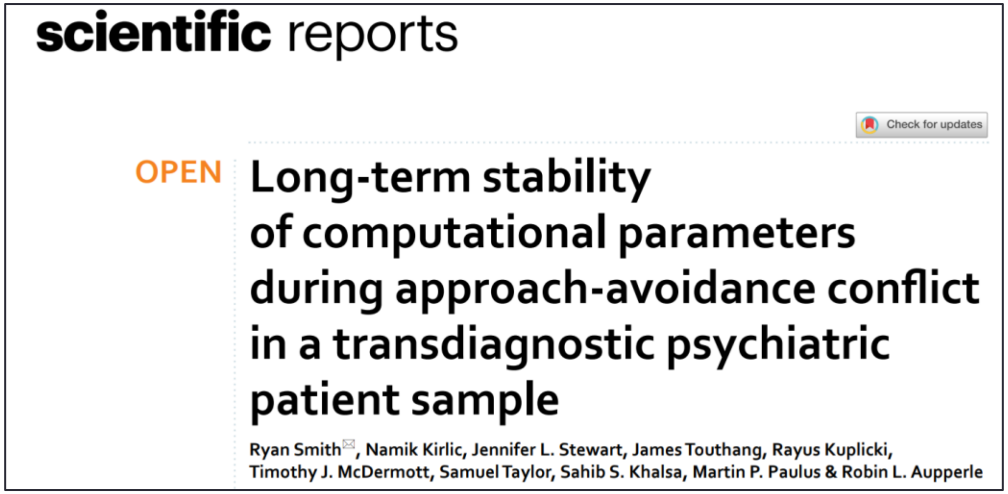
Approach-avoidance conflict in depression, anxiety, and substance use


Journal of Psychiatry & Neuroscience
Submit a Manuscript | Email Alerts | About JPN

J Psychiatry Neurosci. 2021 Jan; 46(1): E74–E87. PMCID: PMC7955838
Prepublished online 2020 Oct 29. doi: 10.1503/jpn.200032 PMID: 33119490

Greater decision uncertainty characterizes a transdiagnostic patient sample during approach-avoidance conflict: a computational modelling approach

Ryan Smith, PhD, Namik Kirlic, PhD, Jennifer L. Stewart, PhD, James Touthang, BS, Rayus Kuplicki, PhD, Sahib S. Khalsa, MD, PhD, Justin Feinstein, PhD, Martin P. Paulus, MD, and Robin L. Aupperle, PhD


scientific reports

Check for updates

OPEN **Long-term stability of computational parameters during approach-avoidance conflict in a transdiagnostic psychiatric patient sample**

Ryan Smith, Namik Kirlic, Jennifer L. Stewart, James Touthang, Rayus Kuplicki, Timothy J. McDermott, Samuel Taylor, Sahib S. Khalsa, Martin P. Paulus & Robin L. Aupperle

Empirical Studies

- **Testing Neuroscientific Predictions**:

Explaining ERPs in response to interoceptive stimuli

ELSEVIER

Biological Psychology

Volume 164, September 2021, 108152

Gut inference: A computational modelling approach

Ryan Smith [a], Ahmad Mayeli [a], Samuel Taylor [a], Obada Al Zoubi [a], Jessyca Naegele [a], Sahib S. Khalsa [a, b]

Dopamine and changes in decision uncertainty

The Dopaminergic Midbrain Encodes the Expected Certainty about Desired Outcomes

Philipp Schwartenbeck, Thomas H. B. FitzGerald, Christoph Mathys, Ray Dolan, Karl Friston   Author Notes

Tutorial

- Most of this talk will draw from our recent tutorial:

Smith, R., Friston, K., & Whyte, C. (2021). **A Step-by-Step Tutorial on Active Inference and its Application to Empirical Data**. *PsyArxiv*. https://doi.org/10.31234/osf.io/b4jm6

Karl Friston

Christopher Whyte

What is active inference?

- Term was initially used to refer to a theory of **predictive motor control** (e.g., Adams et al., 2013)
  - An extension of predictive coding principles
- Currently more often used to refer to a **separate theory of predictive decision-making**
  - (What we are covering in this lecture)
- **Both theories are grounded in Bayesian inference**
  - Deciding **what** to do
  - Controlling the body to **enact that decision**

What is active inference?

- Based on **Partially Observable Markov Decision Processes (POMDPs)**
  - Like normal MDPs, but with **observations** that are **separate from states**
- This means the agent **does not have full knowledge** of the environment
- It must **infer** the 'hidden' states of the world based on **incomplete information** contained in observations

What is active inference?

- The basic idea is that agents are **not simply passive Bayesian observers**

- Instead, agents also *actively* infer the **probability of future observations** given the **different possible actions** they might choose
  - "If I go inside, then I **predict I will feel warm**"

- Combines these predictions with **preferences**
  - "I **want** to feel warm, therefore I will go inside"

- Agents also select actions to **gather observations** that will **aid in precise inference**
  - "I'm going to **turn on my flashlight** because I **predict** this will help me **figure out** how to get inside"

Notation

Bayes' theorem

$$p(s|o,\pi) = \frac{p(o|s,\pi)p(s|\pi)}{p(o|\pi)}$$

- o = observations (also called outcomes)

- s = **hidden** states
  - hidden because they **must be inferred** from observations

- π = policies (possible **sequences of actions**)
  - u = actions

- We **evaluate everything under policies**, because the goal is to choose a policy

- $p(s|\pi)$ = 'prior' belief
  - The **state** I expect to be in if I choose a given **policy**

- $p(o|s,\pi)$ = 'likelihood' mapping
  - The **observations I expect** if I am in this **state** and I choose this **policy**

- $p(s|o,\pi)$ = 'posterior' belief
  - The **state** I infer that I am in **after** making this **observation** if I am following this **policy**
  - This is what we usually want to **figure out**

Notation

Bayes' theorem

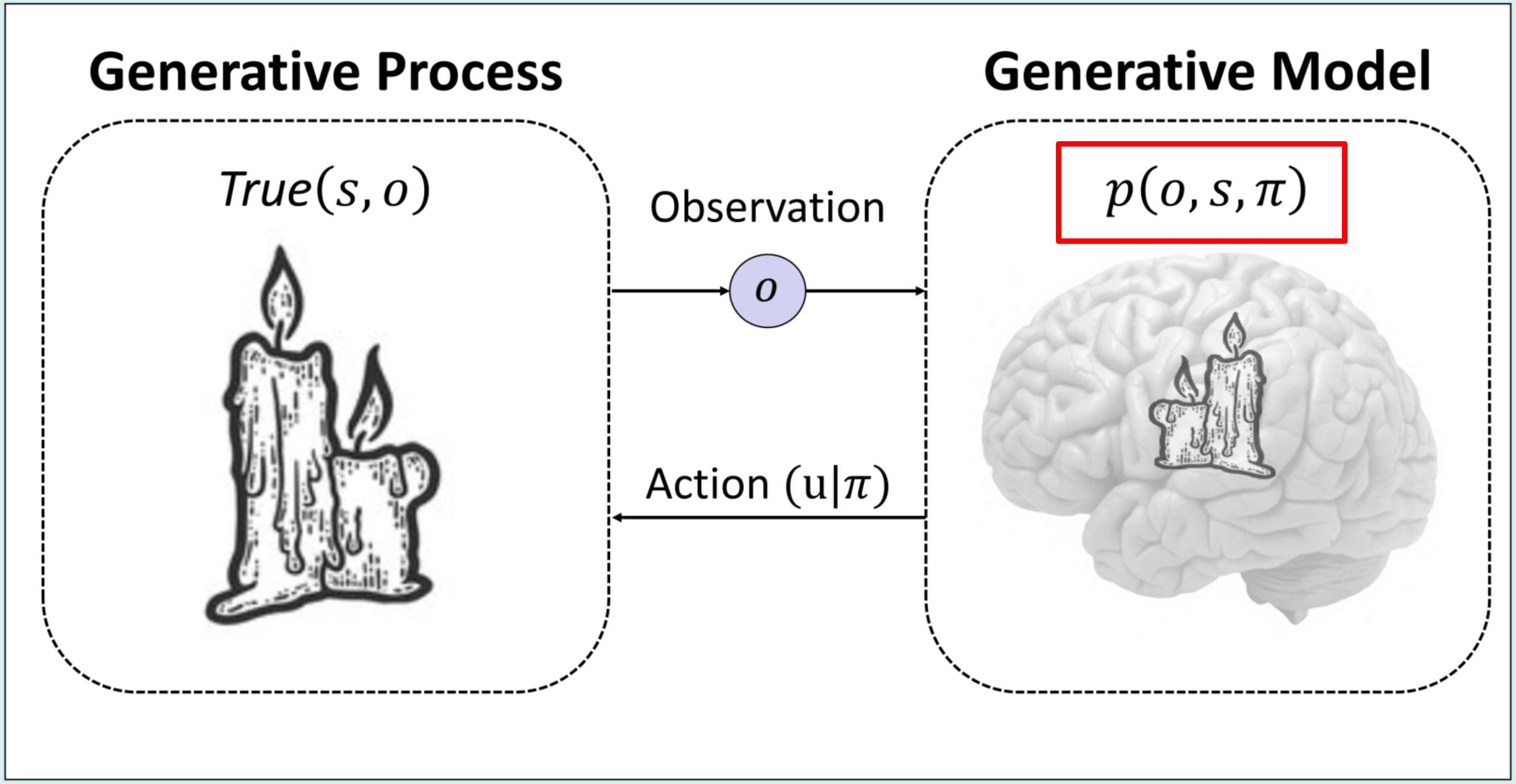
$$p(s|o,\pi) = \frac{p(o|s,\pi)p(s|\pi)}{p(o|\pi)}$$

- $p(o|\pi)$ = 'predictive posterior' over observations
  - The **observations** I expect if I choose this **policy**
- $p(o|C)$ = 'prior preference' distribution
  - The **observations** I find **rewarding**
  - Unique element in active inference – uses something **taking the form of a probability distribution** to encode reward
- q = denotes **approximate** distributions
  - Because **exact** Bayesian inference is often **intractable**
  - $q(s|\pi)$ = *approximate* posterior over states
    - We want to get this to **match $\boldsymbol{p(s|o,\pi)}$ as closely as possible**
  - $q(o|\pi)$ = *approximate* prediction about **what I will observe if I choose this policy**

Notation

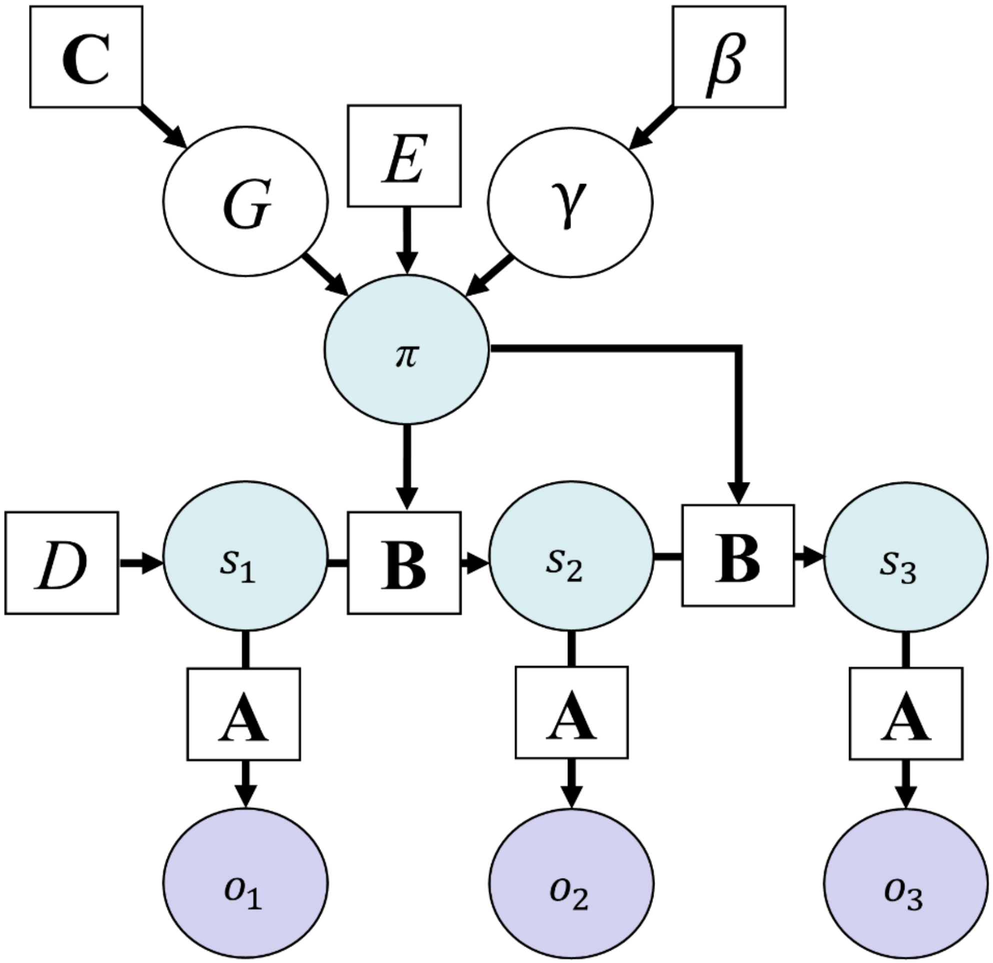
$$D_{KL}[p(x)||p(y)] = \sum_x p(x)\left[\ln\frac{p(x)}{p(y)}\right]$$

- $D_{KL}[p(x)||p(y)]$ – the *Kullback–Leibler (KL) divergence*
  - Larger values indicate that two **distributions are more dissimilar**
  - The closer we can get $q(s|\pi)$ to match $p(s|o,\pi)$, the smaller their D_{KL} will be
- τ (Greek letter tau) – the time **about which** I have a belief
  - I believe I **am** now in my car (s_τ)
  - I believe I **was** in my kitchen 10 min ago ($s_{\tau-1}$)
  - I believe I **will be** at work in 20 min ($s_{\tau+1}$)
- t – the time **at which** a **new observation** is given
  - After turning on a light, I **now** (s_τ) believe that I **was** in the kitchen for the last 5 minutes ($s_{\tau-1}$)
    - (although I didn't know this beforehand)
- So I can **update my belief about all $\boldsymbol{\tau}$s ($\boldsymbol{s_\tau}$)** with each new observation (i.e., at each t)



- *Generative process* – **actual objects/events** generating observations
- *Generative model* – **beliefs** about the generative process and how it will be **affected by action**
    - Probability of **all possible combinations** of observations, states, and policies

Graphical model depiction of active inference



- Here **arrows** denote **conditional dependencies**
- **Circles** correspond to **random variables** that are updated during inference
- **Squares** indicate fixed parameters
- Can be **a bit daunting** for those new to this area
- We will **walk through this step by step**

Task instructions: “On each trial you have the choice of two slot machines. One will pay out 4 dollars, and the other will pay out $0. You can either guess right away or you can ask for a hint about which one is more likely to pay out. However, if you ask for the hint, you can only win 2 dollars.”

Example

Example states and outcomes

Hidden State Factor 1: Context
- Left machine is more likely to pay out
- Right machine is more likely to pay out

Hidden State Factor 2: Choice states
- Start
- Ask for the hint
- Choose left machine
- Choose right machine

Outcome Modality 1: Hint
- No hint
- Left machine is more likely
- Right machine is more likely

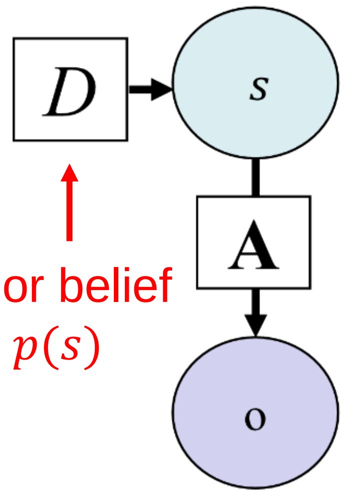
Outcome Modality 2: Outcome
- Start
- Lose
- Win

Static perception

Equation

$$s = \sigma\left(\ln D + \ln \mathbf{A}^{\mathrm{T}} o\right)$$

Equivalent to Bayes theorem
$p(s|o) \propto p(s)p(o|s)$



s = hidden states

Prior belief
$p(s)$

A = $p(o|s)$ = likelihood mapping

o = observation

- σ = a softmax (normalized exponential) function - **transforms a vector into a probability distribution**
  - (non-negative values that sum to 1)

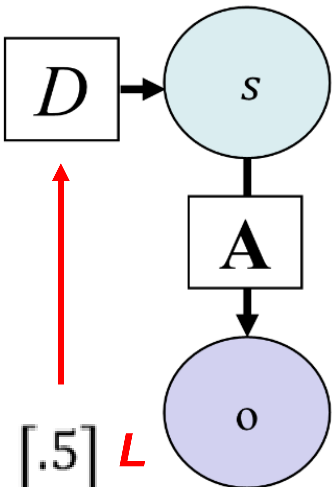
Static perception

Equation

$$\boxed{s = \sigma\left(\ln D + \ln \mathbf{A}^{\mathrm{T}} o\right)}$$

$$s = \begin{bmatrix} .82 \\ .18 \end{bmatrix}$$

Left machine is more likely to be the better one



$$\mathbf{A} = \begin{array}{c} \begin{array}{cc} \mathit{L} & \mathit{R} \end{array} \\ \begin{bmatrix} \boxed{.9 \quad .2} \\ .1 \quad .8 \end{bmatrix} \end{array}$$

p(hint=left machine | left machine is better) = .9

$$o = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Hint = left machine

$$D = \begin{array}{cc} \begin{bmatrix} .5 \\ .5 \end{bmatrix} & \begin{array}{c} \mathit{L} \\ \mathit{R} \end{array} \end{array}$$

Both machines are equally likely

$$\boxed{s = \sigma\left(\ln \begin{bmatrix} .5 \\ .5 \end{bmatrix} + \ln \begin{bmatrix} .9 & .1 \\ .2 & .8 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \sigma\left(\ln \begin{bmatrix} .5 \\ .5 \end{bmatrix} + \ln \begin{bmatrix} .9 \\ .2 \end{bmatrix}\right) = \sigma\left(\ln \begin{bmatrix} .45 \\ .10 \end{bmatrix}\right)}$$

Dynamic perception

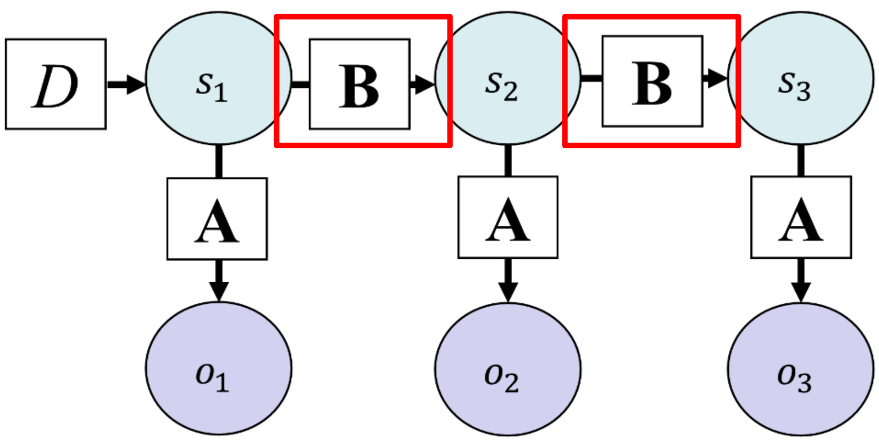
- $\mathbf{B}$ = matrix encoding probability of state transitions, $p(s_{\tau+1}|s_\tau)$
  - Provide prior beliefs for states at times $\tau > 1$

Prior, $p(s)$ Likelihood, $p(o|s)$

Equations

$$s_{\tau=1} = \sigma\left(\frac{1}{2}\boxed{\left(\ln D + \ln(\mathbf{B}_\tau^\dagger\, s_{\tau+1})\right)} + \boxed{\ln \mathbf{A}^\mathrm{T}\, o_\tau}\right)$$

$$s_{\tau>1} = \sigma\left(\frac{1}{2}\boxed{\left(\ln(\mathbf{B}_{\tau-1}\, s_{\tau-1}) + \ln(\mathbf{B}_\tau^\dagger\, s_{\tau+1})\right)} + \boxed{\ln \mathbf{A}^\mathrm{T}\, o_\tau}\right)$$



† indicates a transpose and then normalizing columns

Dynamic perception

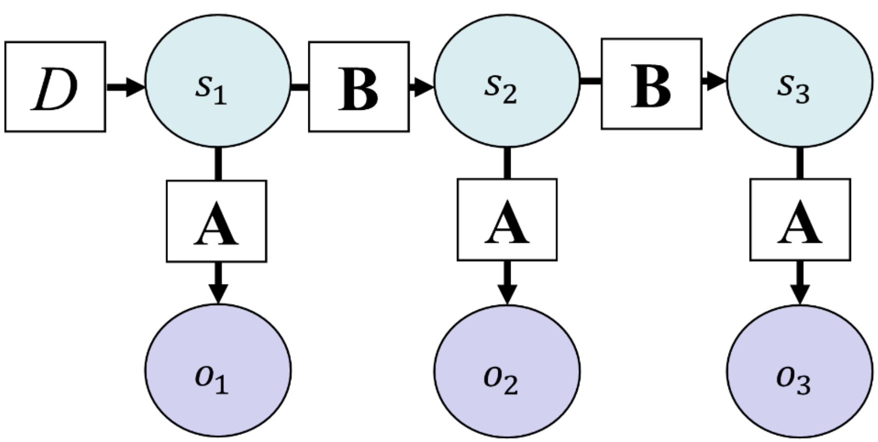
- Allows you to update beliefs about earlier time points when getting observations at later time points (τ vs. t)

Prior from the future

Equations

$$s_{\tau=1} = \sigma\left(\frac{1}{2}\left(\ln D + \boxed{\ln(\mathbf{B}_\tau^\dagger\, s_{\tau+1})}\right) + \ln \mathbf{A}^\mathrm{T}\, o_\tau\right)$$

$$s_{\tau>1} = \sigma\left(\frac{1}{2}\left(\ln(\mathbf{B}_{\tau-1}\, s_{\tau-1}) + \boxed{\ln(\mathbf{B}_\tau^\dagger\, s_{\tau+1})}\right) + \ln \mathbf{A}^\mathrm{T}\, o_\tau\right)$$

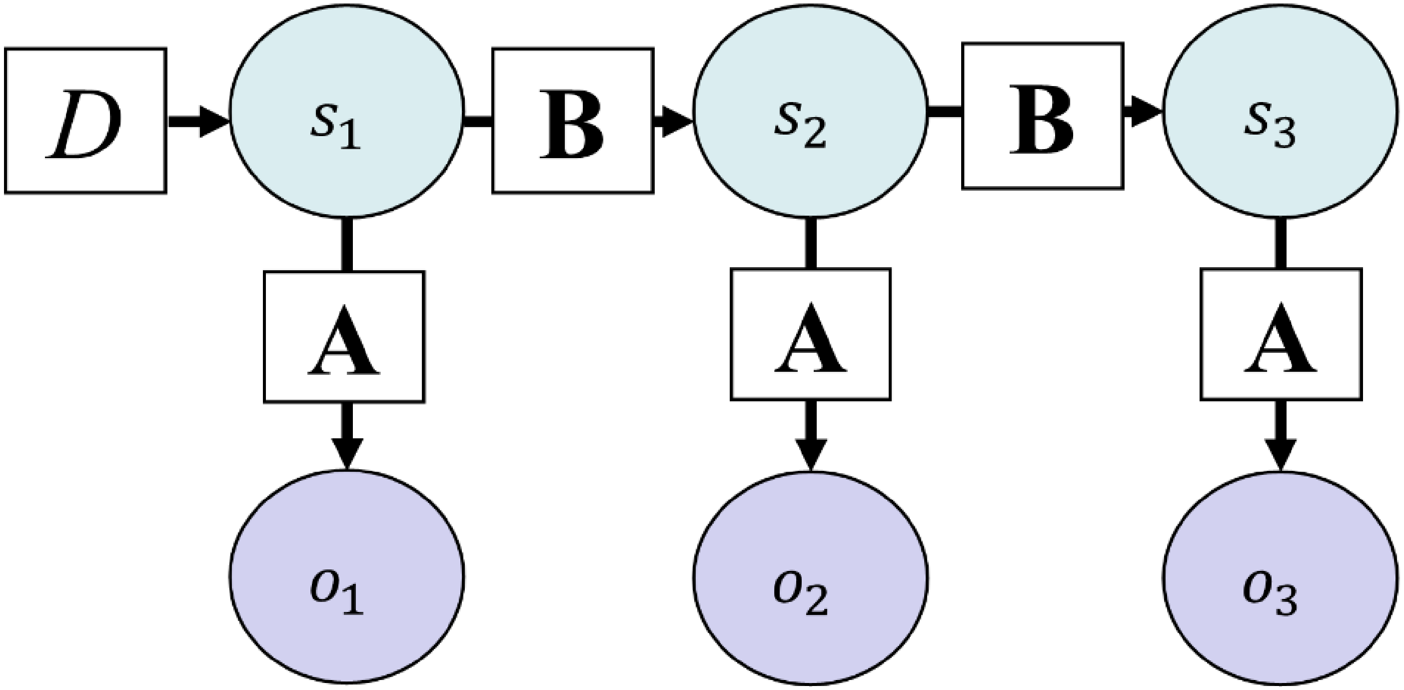


† indicates a transpose and then normalizing columns

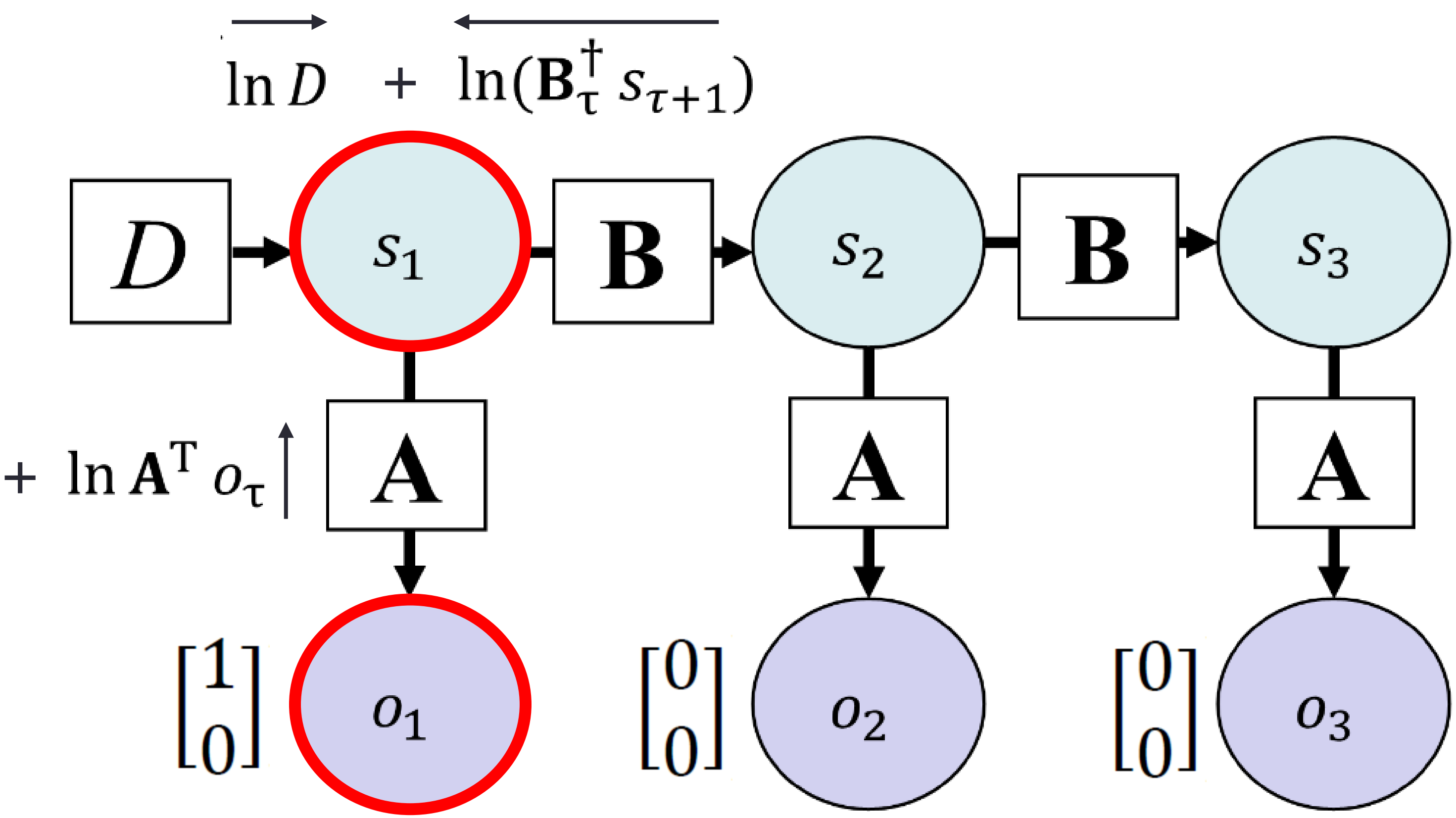
Parr, Thomas, et al. "Neuronal message passing using Mean-field, Bethe, and Marginal approximations." *Scientific reports* 9.1 (2019): 1-18.

Variational (Marginal) Message passing

- These equations implement **variational message passing**
  - An efficient way to do *approximate* **posterior inference on a graph**
  - Based on **minimizing** a quantity called **variational free energy (F)**
    - Will **first illustrate how this works** and **then introduce F formally**

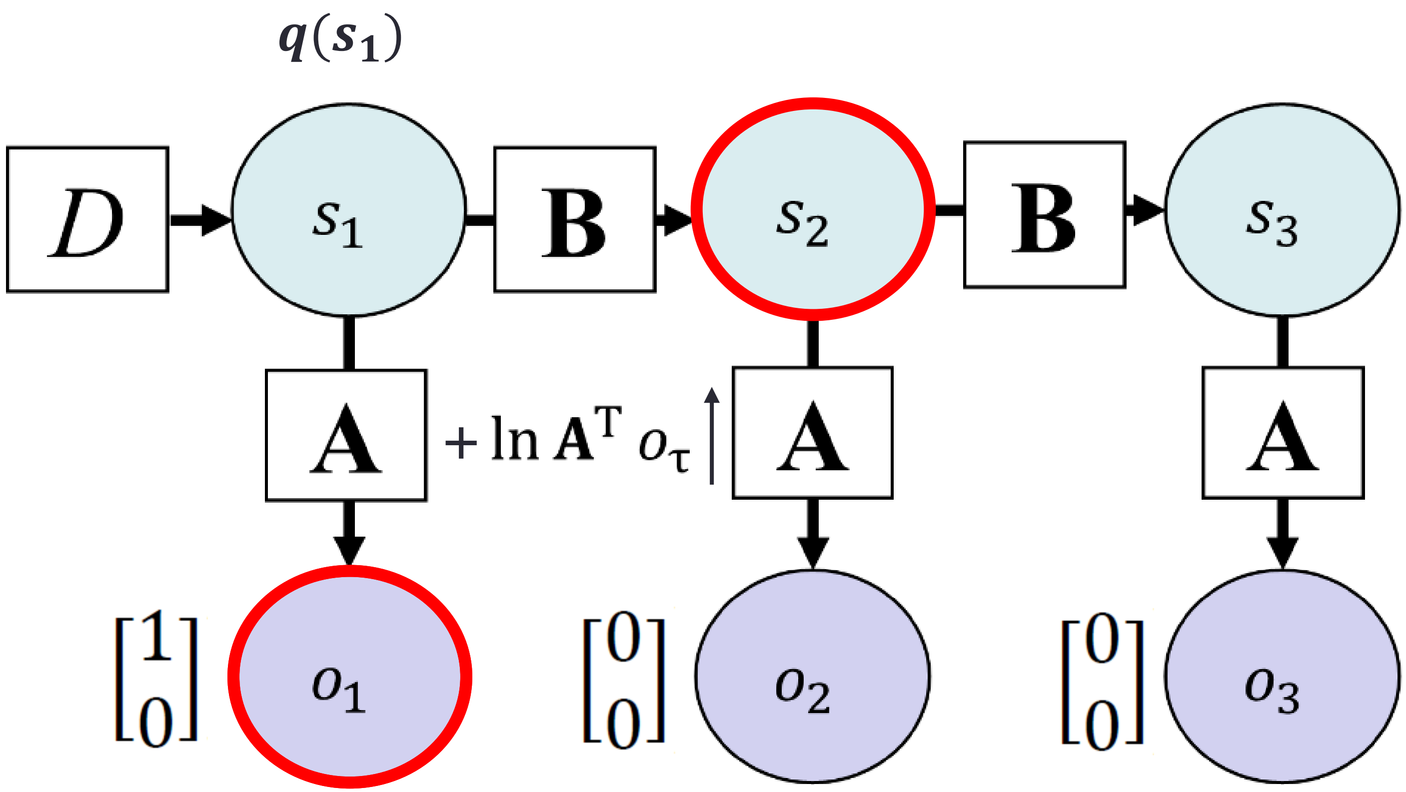


Variational Message passing

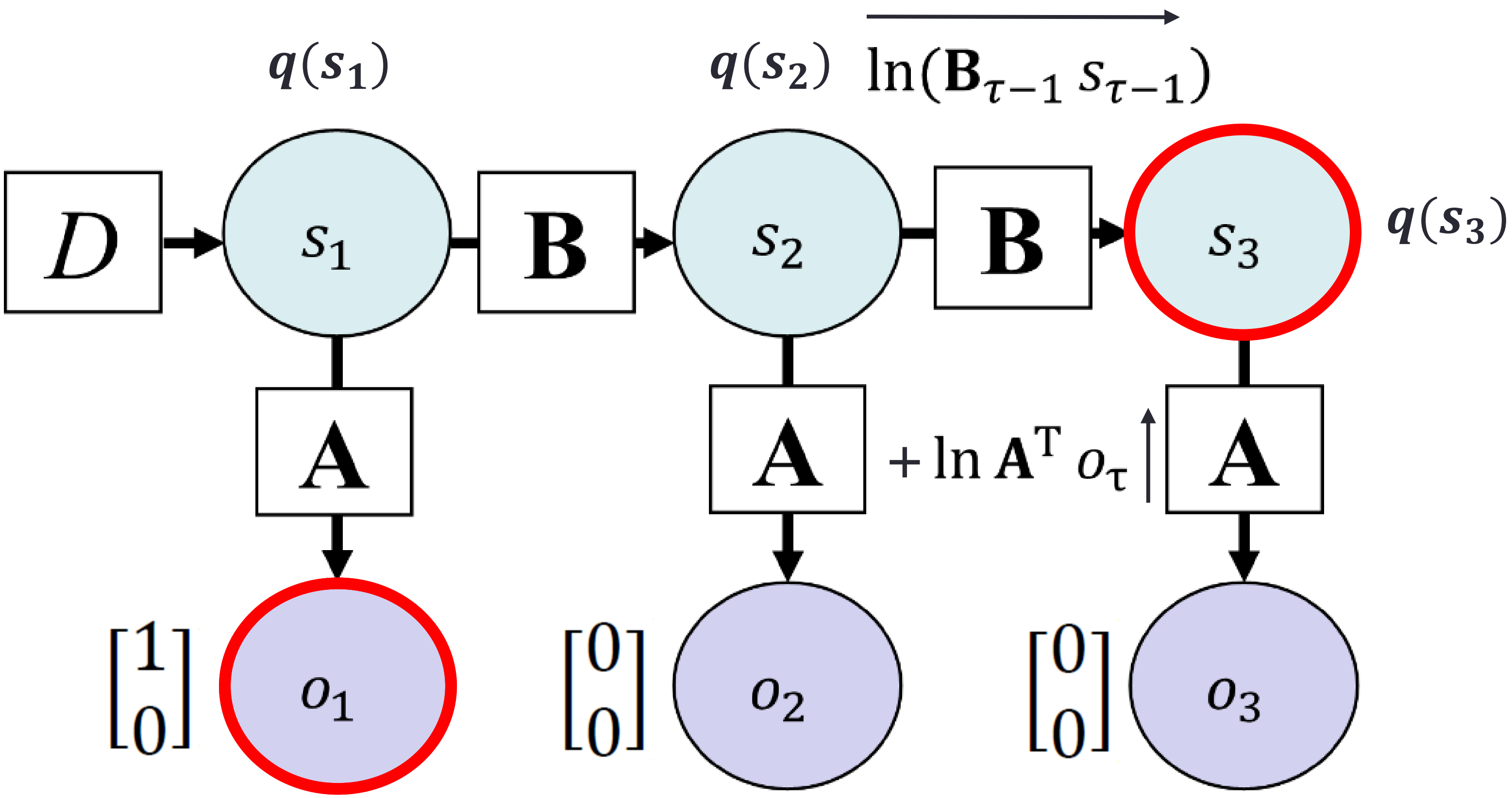


Variational Message passing

$$\overrightarrow{\ln(\mathbf{B}_{\tau-1}\, s_{\tau-1})} \;+\; \overleftarrow{\ln(\mathbf{B}_{\tau}^{\dagger}\, s_{\tau+1})}$$

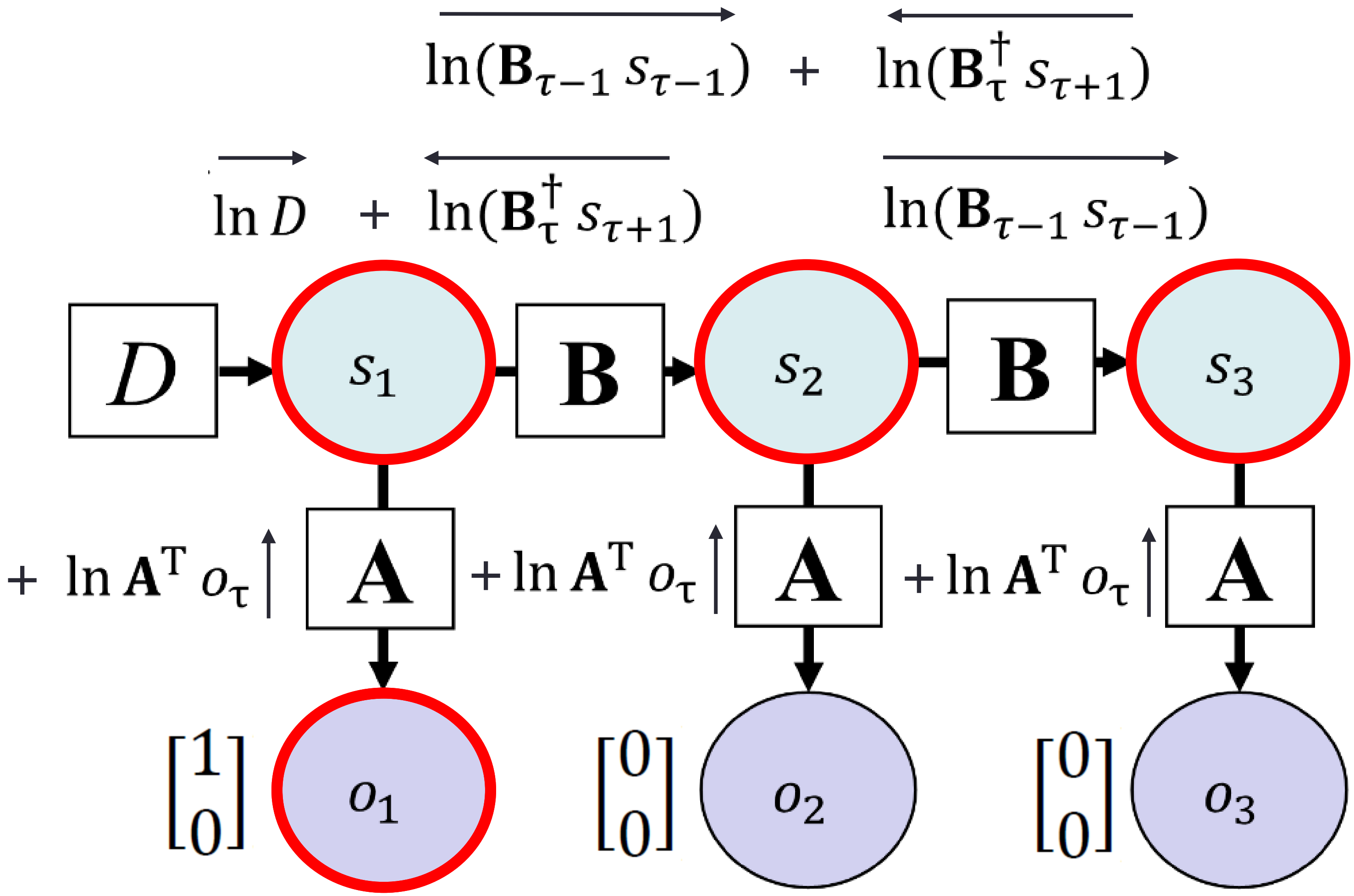


Variational Message passing



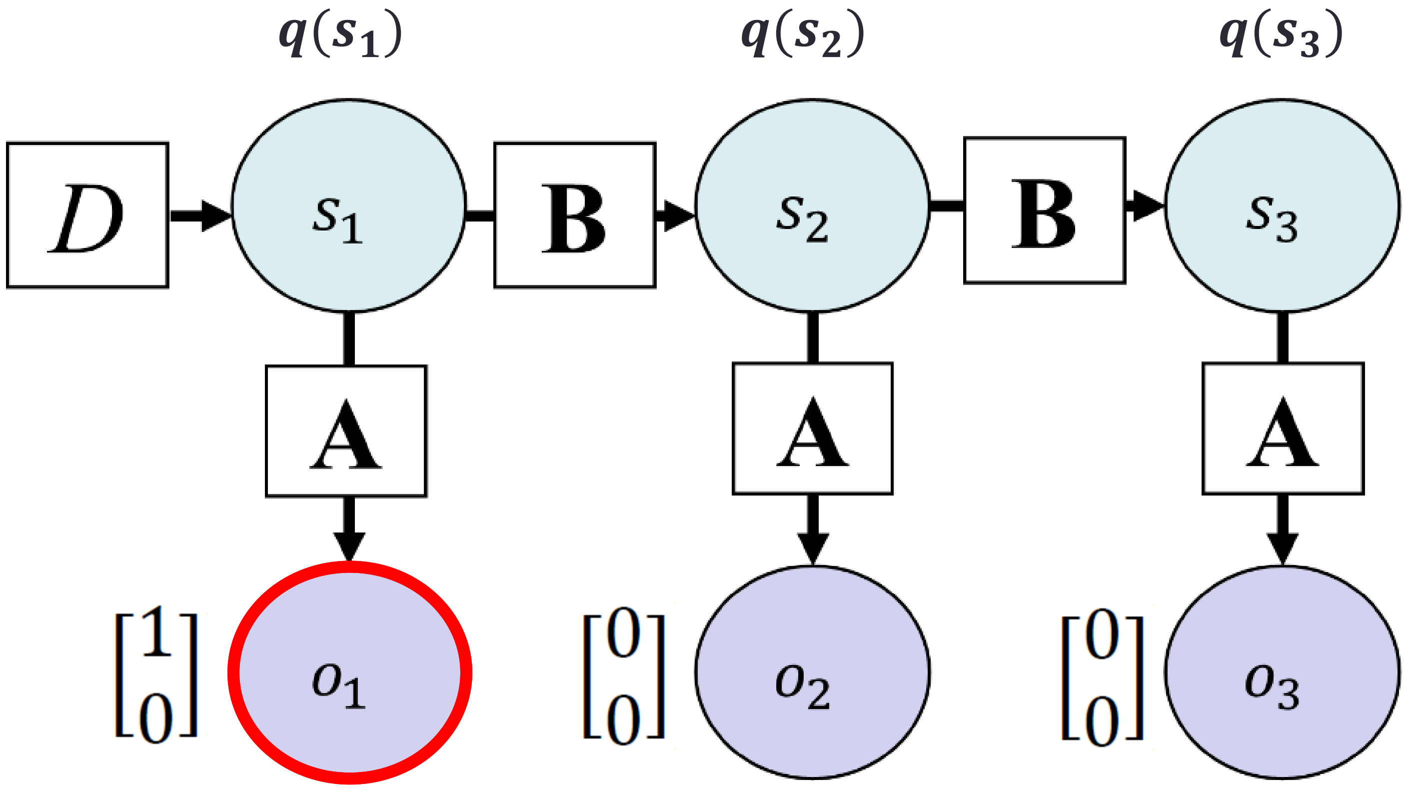
Variational Message passing

- **Iterate** this process until convergence to a stable s for each τ



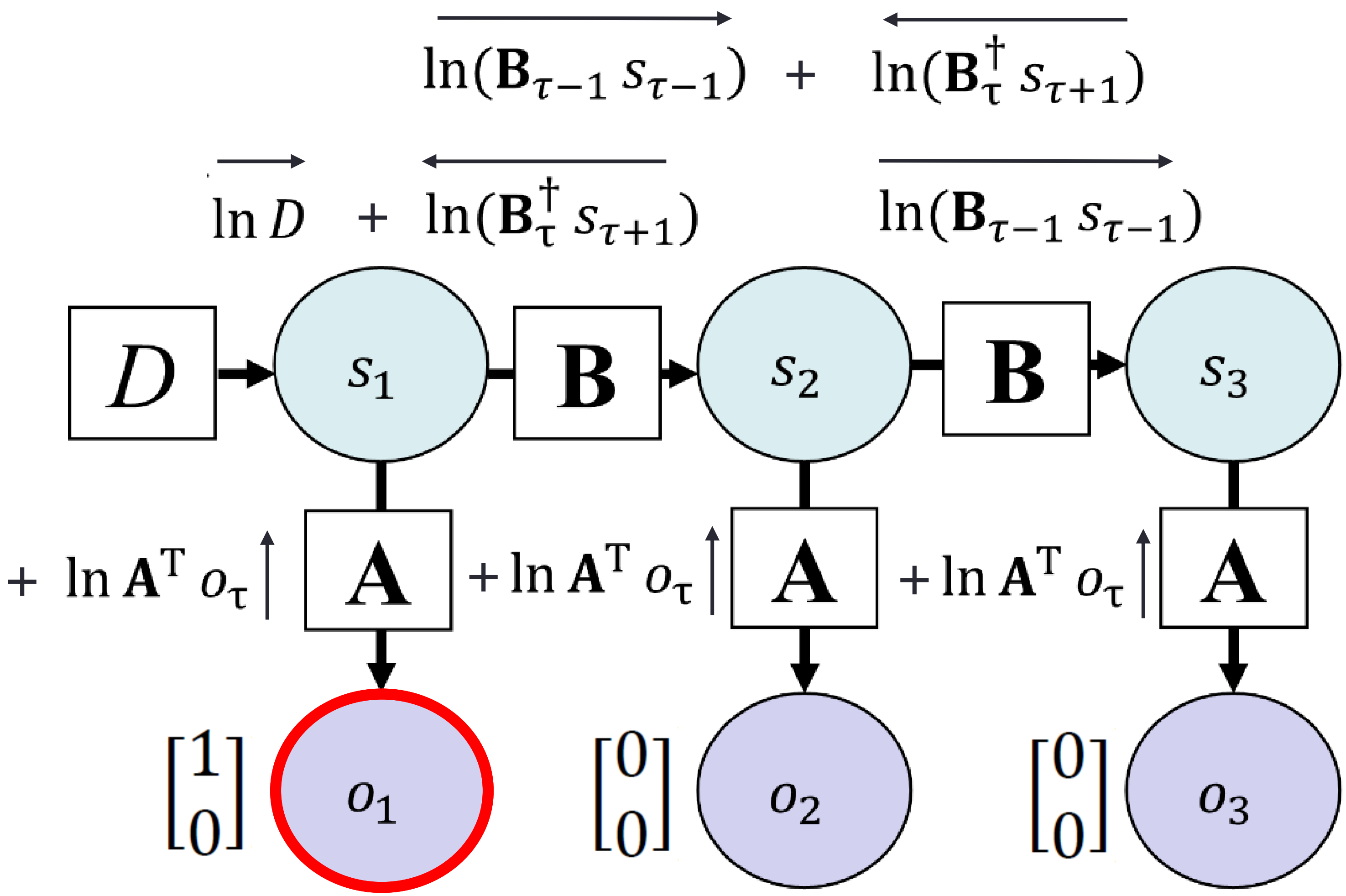
Variational Message passing

- **Iterate** this process until convergence to a stable s for each τ



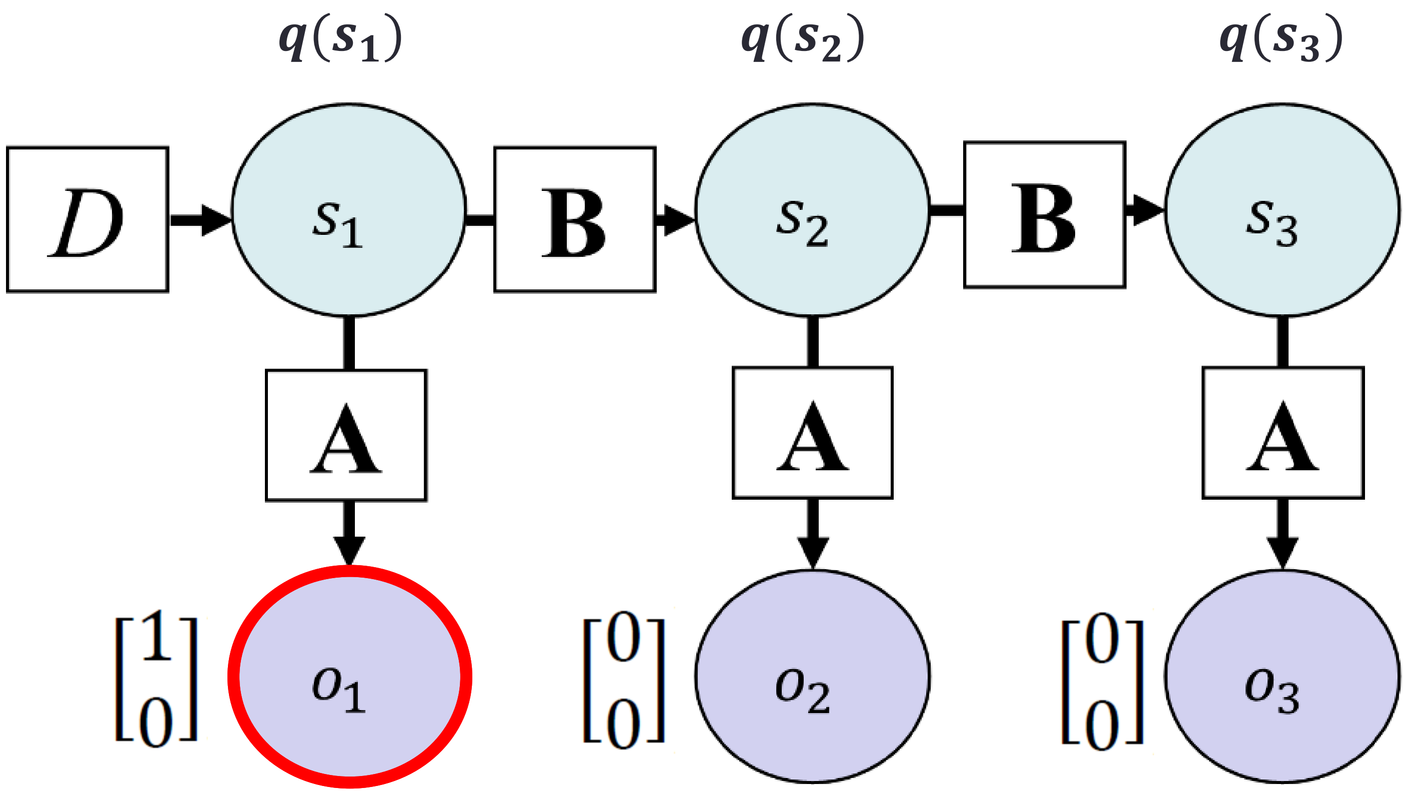
Variational Message passing

- **Iterate** this process until convergence to a stable s for each τ



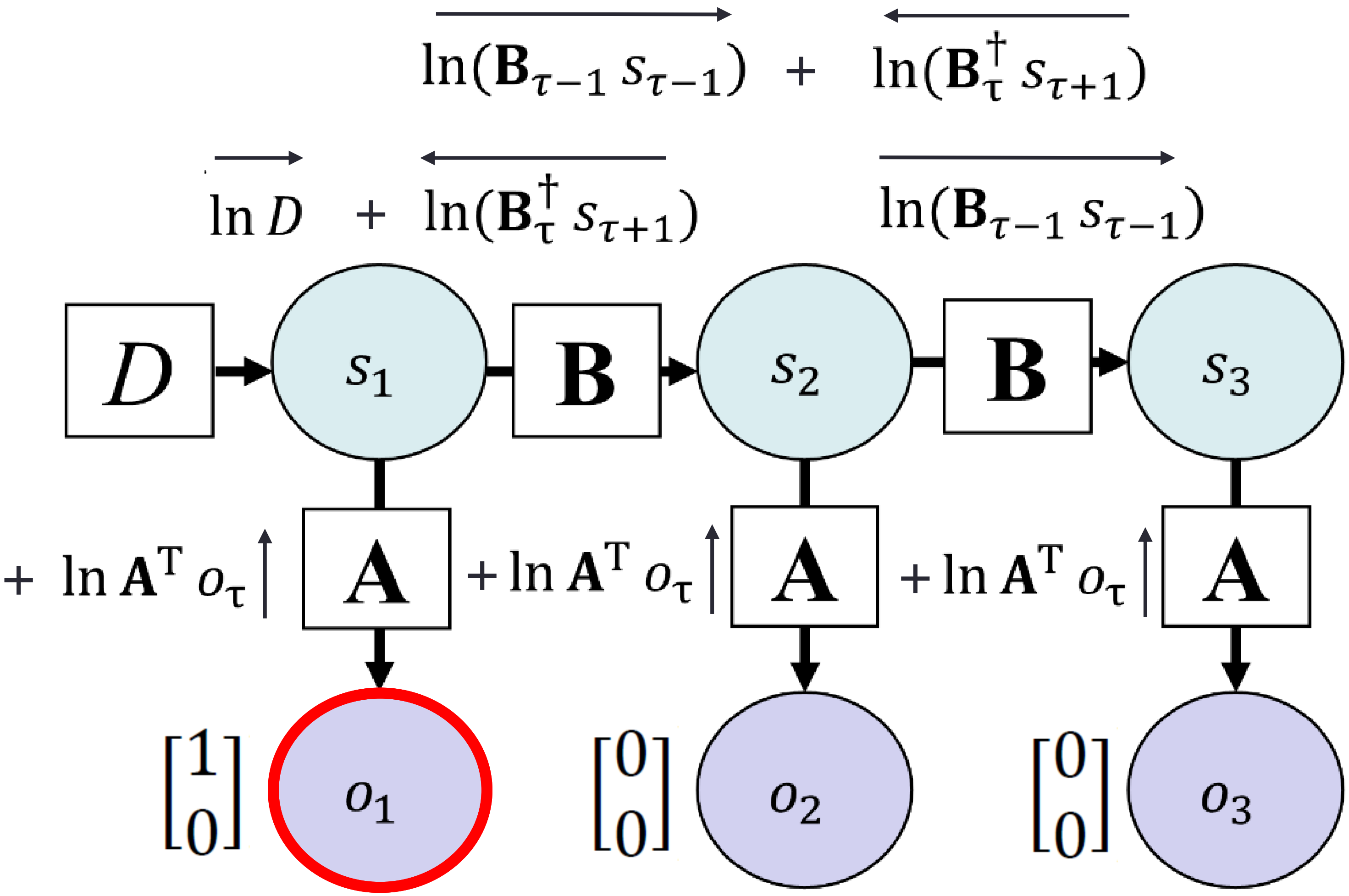
Variational Message passing

- **Iterate** this process until convergence to a stable s for each τ



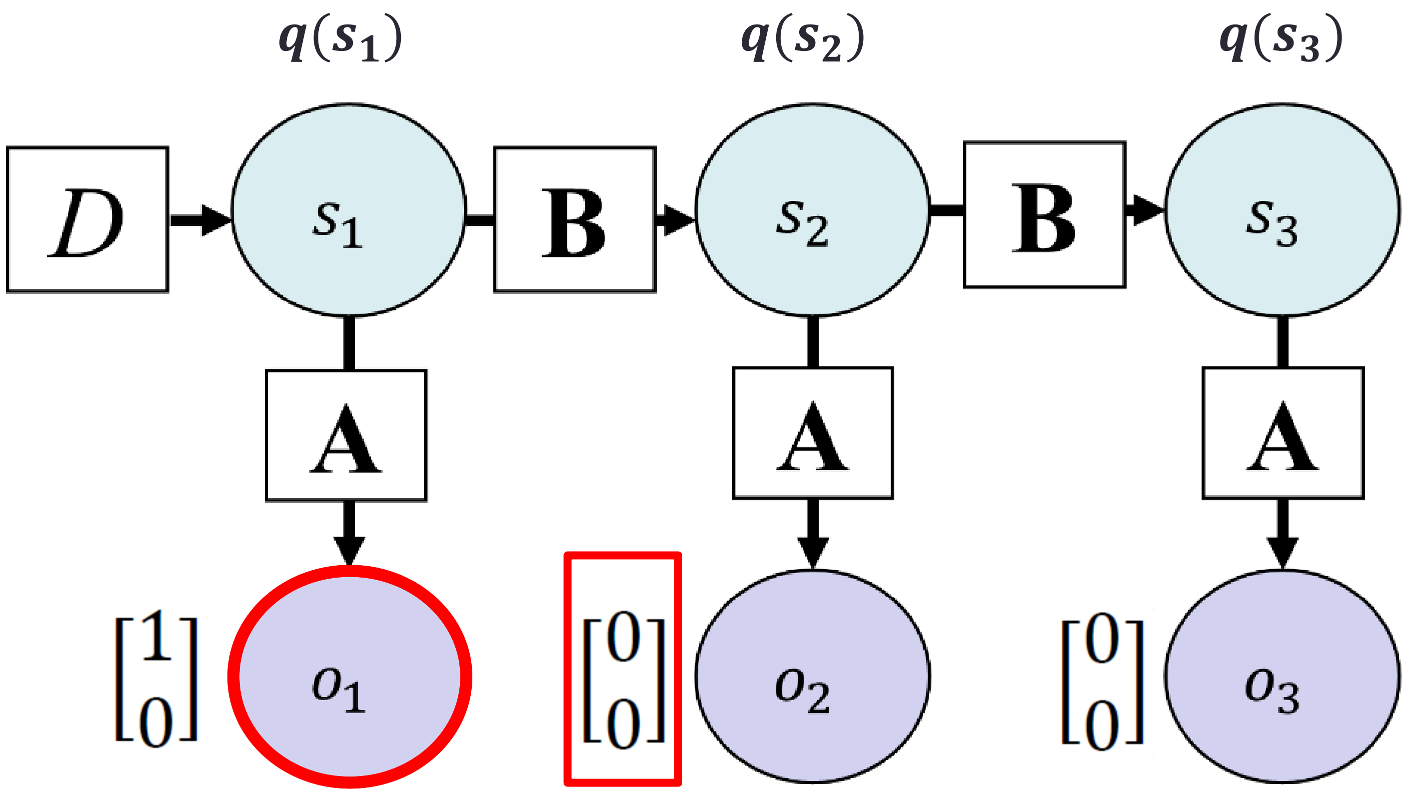
Variational Message passing

- **Iterate** this process until convergence to a stable s for each τ



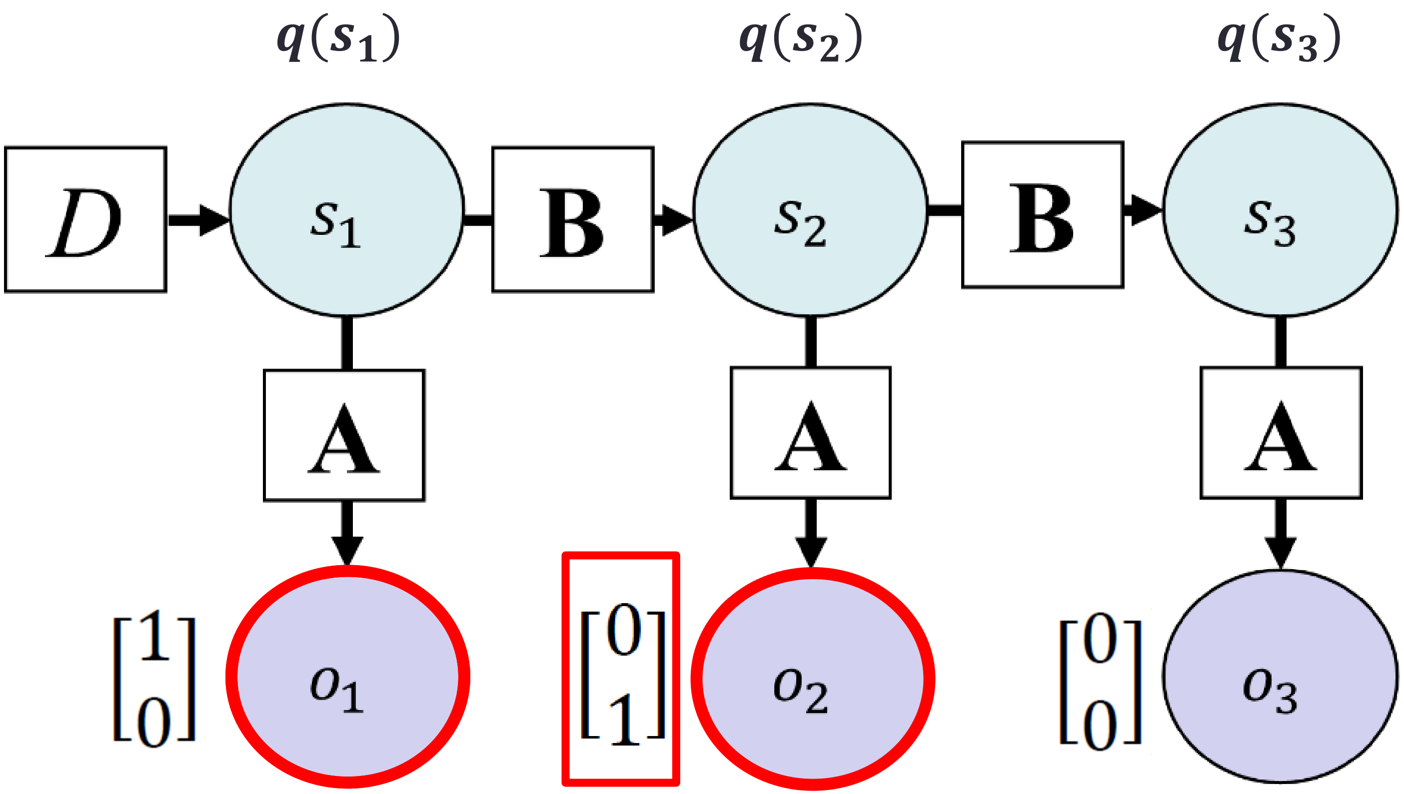
Variational Message passing

- **Let’s now assume beliefs over states have converged**



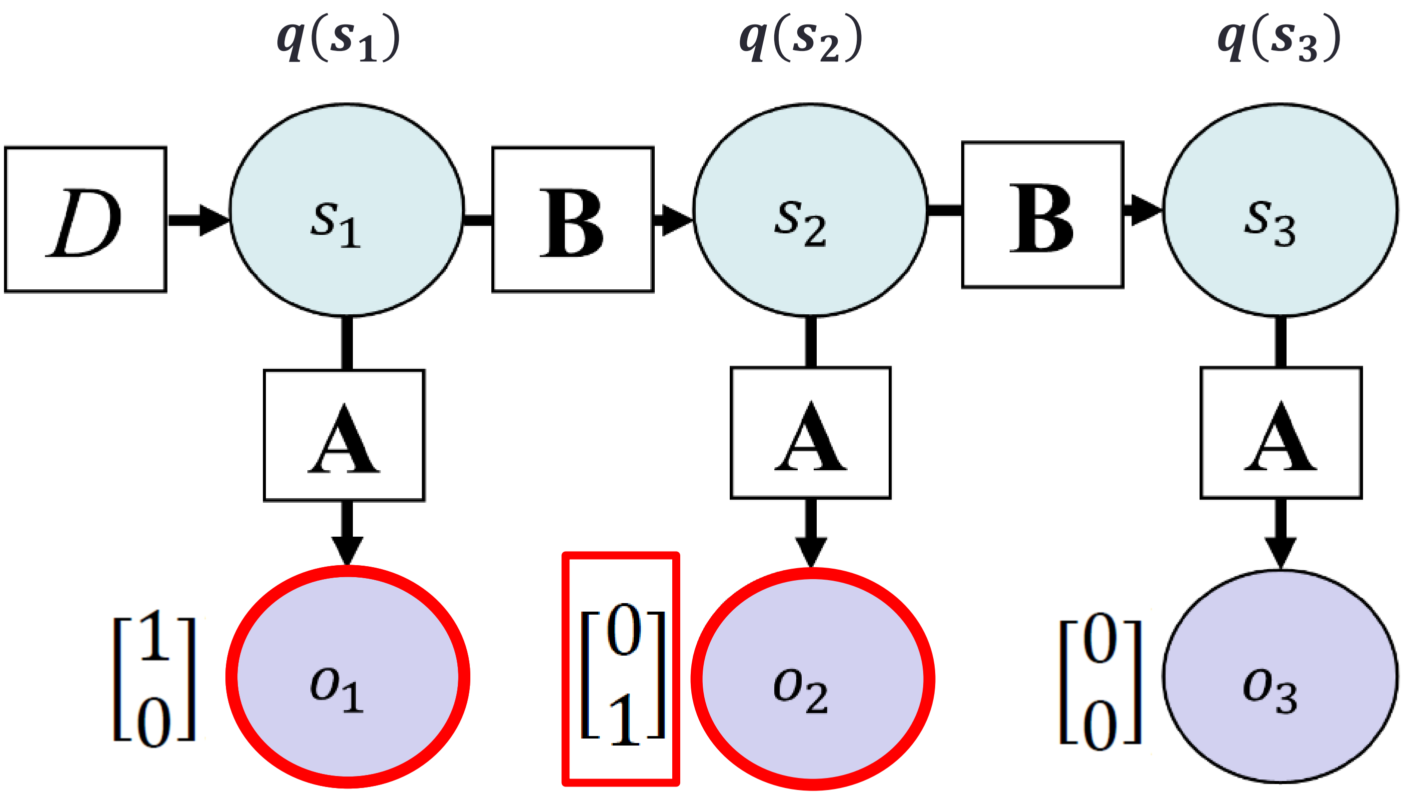
Variational Message passing

- **Add next observation**
  - **Repeat** iterative message passing until you **again reach convergence**, etc.



Variational Message passing

- Also represents a possible way to think about the **messages passed between neurons** in the brain
  - Message passing can be formulated in terms of **minimizing prediction error**



(Note that the prediction error version I showed here was based on a model where predictions are conditioned on policies, but nothing else changes if policy subscripts are removed)

Neural process theory

$$s_{\tau>1} = \sigma\left(\frac{1}{2}\left(\ln(\mathbf{B}_{\tau-1}\, s_{\tau-1}) + \ln(\mathbf{B}_{\tau}^{\dagger}\, s_{\tau+1})\right) + \ln \mathbf{A}^{\mathrm{T}}\, o_\tau\right)$$

State prediction error

Approximate posterior over states

$$\varepsilon_{\pi,\tau} \leftarrow \overbrace{\frac{1}{2}\left(\ln\left(\mathbf{B}_{\pi,\tau-1} s_{\pi,\tau-1}\right) + \ln\left(\mathbf{B}_{\pi,\tau}^{\dagger} s_{\pi,\tau+1}\right)\right) + \ln \mathbf{A}^{\mathrm{T}} o_\tau}^{\text{Generative model after observation}} - \ln s_{\pi,\tau}$$

$$v_{\pi,\tau} \leftarrow v_{\pi,\tau} + \varepsilon_{\pi,\tau} \qquad v_{\pi,\tau} = \ln s_{\pi,\tau}$$

Membrane voltage

$$s_{\pi,\tau} \leftarrow \sigma(v_{\pi,\tau})$$

Normalized Firing rate

$$-\boldsymbol{\nabla}_{s_{\pi,\tau}} \boldsymbol{F}_{\pi} = \boldsymbol{\varepsilon}_{\pi,\tau}$$

Task instructions: "On each trial you have the choice of two slot machines. One will pay out 4 dollars, and the other will pay out $0. You can either guess right away or you can ask for a hint about which one is more likely to pay out. However, if you ask for the hint, you can only win 2 dollars."

Example

$$p(s_{\tau+1}|s_\tau) =$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}$$

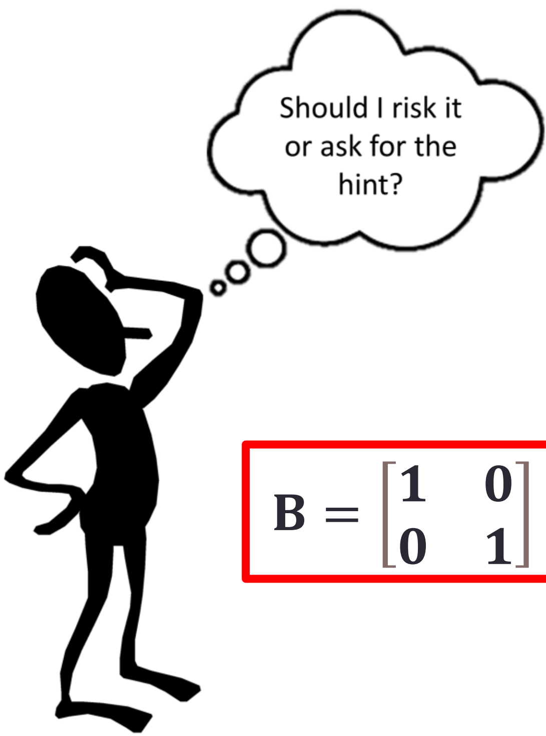
If the left machine is better **at time point 2**

it will **still be better at time point 3**

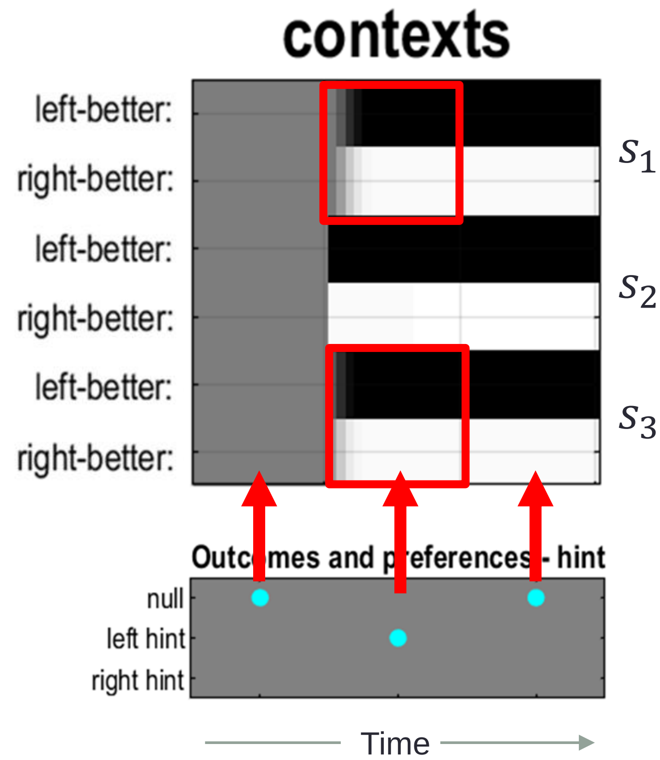
with **probability = 1**

Task instructions: "On each trial you have the choice of two slot machines. One will pay out 4 dollars, and the other will pay out $0. You can either guess right away or you can ask for a hint about which one is more likely to pay out. However, if you ask for the hint, you can only win 2 dollars."

Example



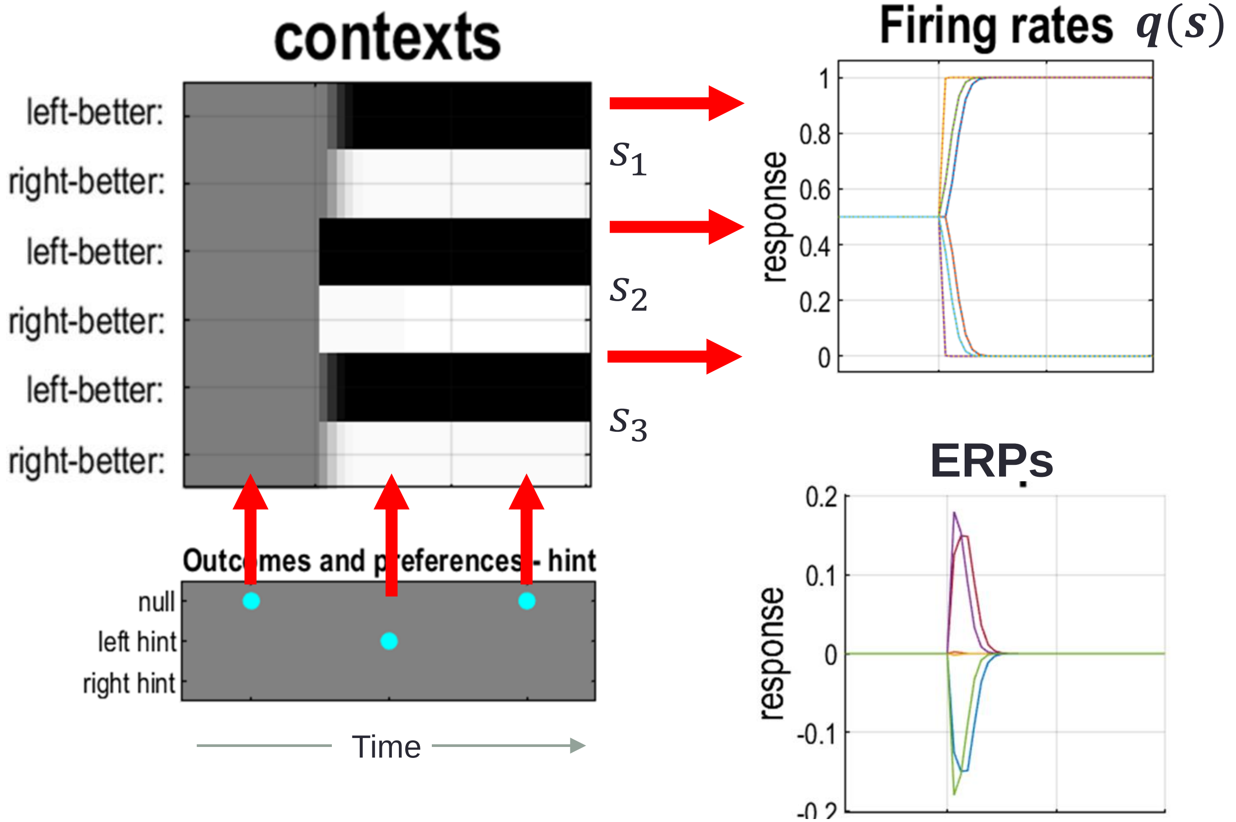
$$\boxed{B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}$$



Darker = higher probability
Cyan dot = true state/outcome/action

Task instructions: "On each trial you have the choice of two slot machines. One will pay out 4 dollars, and the other will pay out $0. You can either guess right away or you can ask for a hint about which one is more likely to pay out. However, if you ask for the hint, you can only win 2 dollars."

Neural process theory



Rate of change in $q(s)$

Task instructions: "On each trial you have the choice of two slot machines. One will pay out 4 dollars, and the other will pay out $0. You can either guess right away or you can ask for a hint about which one is more likely to pay out. However, if you ask for the hint, you can only win 2 dollars."

Example

$$p(s_{\tau+1}|s_\tau) =$$

$$\mathbf{B} = \begin{bmatrix} .7 & .3 \\ .3 & .7 \end{bmatrix}$$

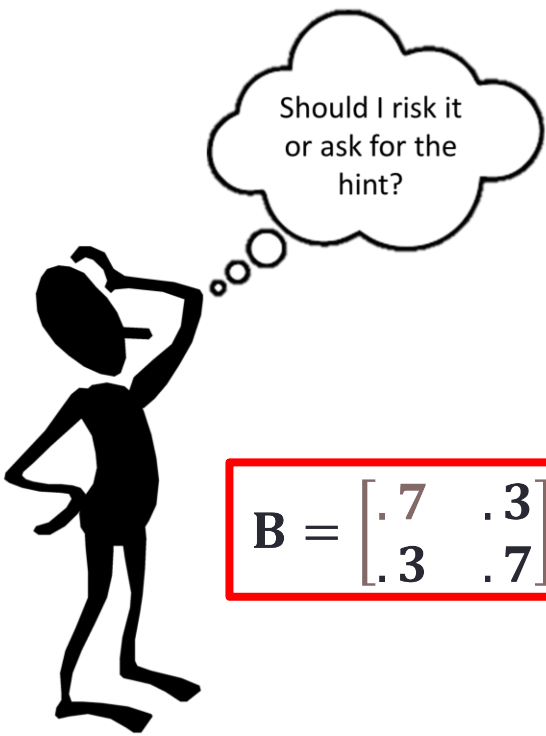
If the left machine is better **at time point 2**

it will **still be better at time point 3**

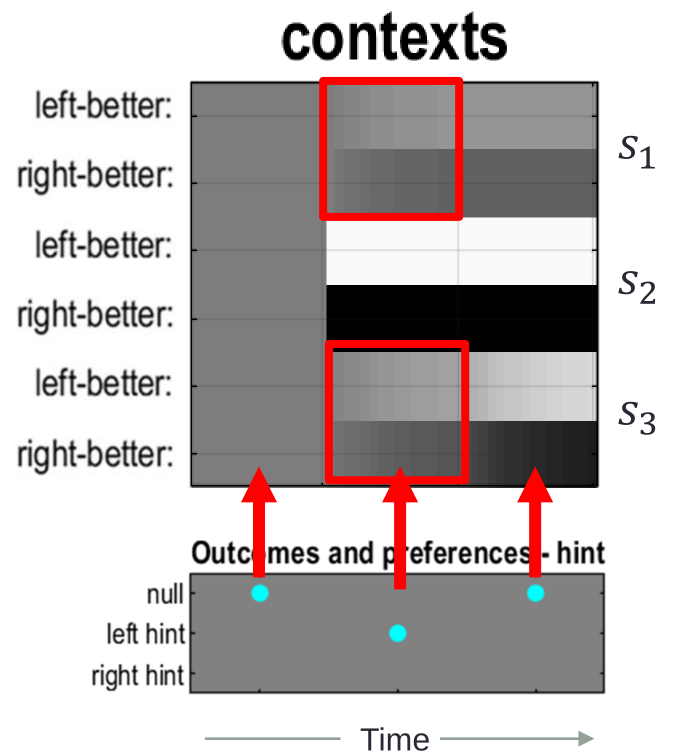
with **probability = .7**

Task instructions: "On each trial you have the choice of two slot machines. One will pay out 4 dollars, and the other will pay out \$0. You can either guess right away or you can ask for a hint about which one is more likely to pay out. However, if you ask for the hint, you can only win 2 dollars."

Example



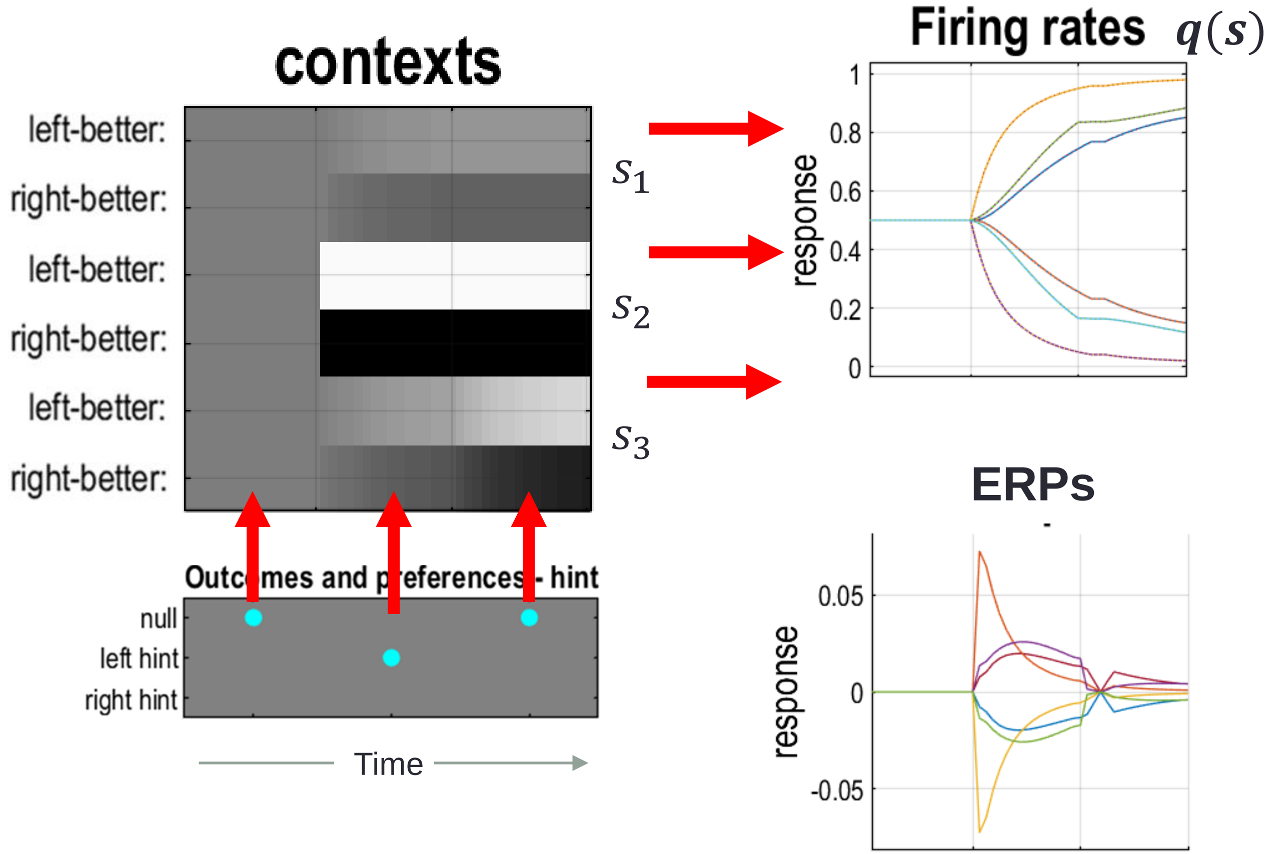
$$\boxed{\mathbf{B} = \begin{bmatrix} .7 & .3 \\ .3 & .7 \end{bmatrix}}$$



Darker = higher probability
Cyan dot = true state/outcome/action

Task instructions: “On each trial you have the choice of two slot machines. One will pay out 4 dollars, and the other will pay out $0. You can either guess right away or you can ask for a hint about which one is more likely to pay out. However, if you ask for the hint, you can only win 2 dollars.”

Neural process theory



Rate of change in $q(s)$

Evidence for predicted ERPs

- A few papers have now shown that ERPs predicted based on this theory can **reproduce well known ERP findings** in the EEG literature.

Working memory, attention, and salience in active inference

Thomas Parr & Karl J Friston

Scientific Reports **7**, Article number: 14678 (2017) | Cite this article

ELSEVIER

Biological Psychology
Volume 164, September 2021, 108152

Gut inference: A computational modelling approach

Ryan Smith [a], Ahmad Mayeli [a], Samuel Taylor [a], Obada Al Zoubi [a], Jessyca Naegele [a], Sahib S. Khalsa [a, b]

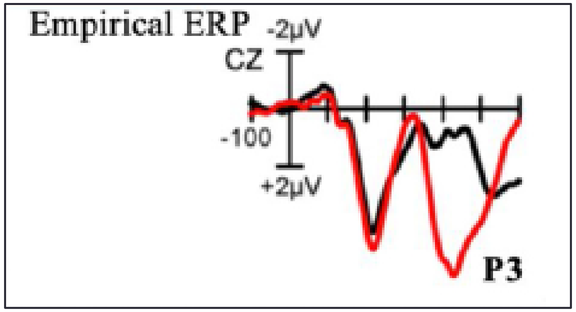
ELSEVIER

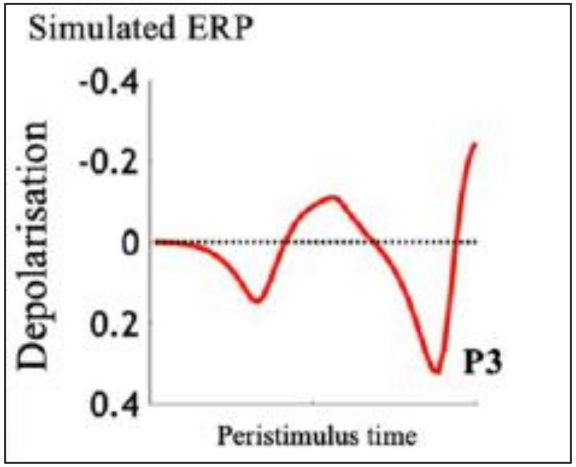
Progress in Neurobiology
Volume 199, April 2021, 101918

Original Research Article

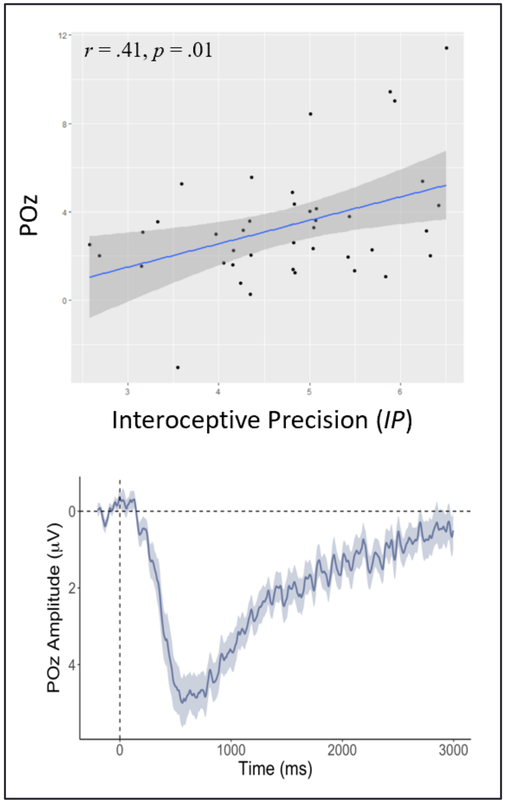
The predictive global neuronal workspace: A formal active inference model of visual consciousness

Christopher J. Whyte [a], Ryan Smith [b]

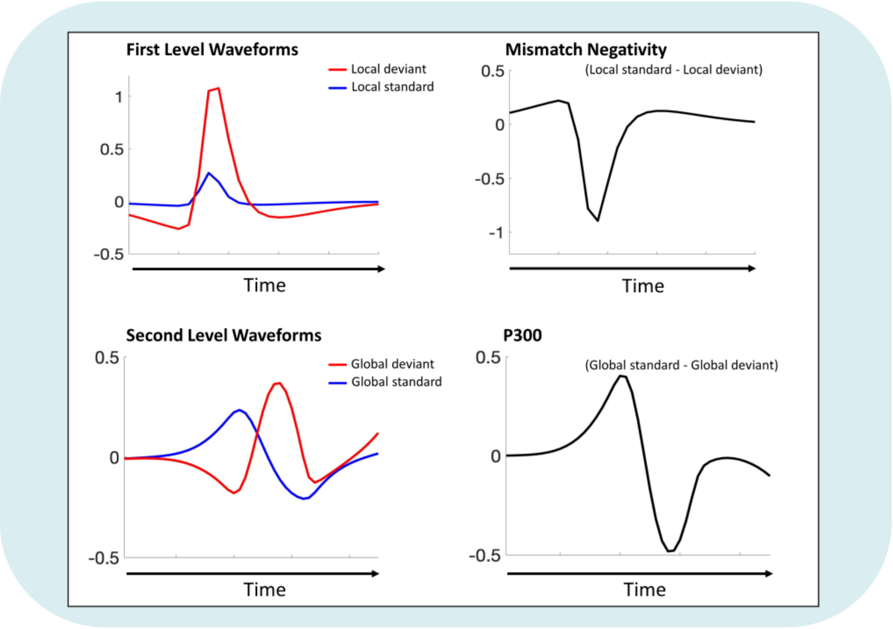




Inattentional blindness task (P300)



Interoception (LPP)



Oddball task (mismatch negativity)

Variational free energy

- Start with generative model you know, $m = p(o, s) = p(o|s)p(s)$
- Introduce an **initially arbitrary distribution $q(s)$ that will approximate the true posterior** we want, $p(s|o)$

$$F = \boxed{\mathrm{E}_{q(s)}[\ln \frac{q(s)}{p(s|o)}]} - \boxed{\ln p(o)}$$

Difference between **approximate** and **true posterior**

"How **surprising** is this observation under my model?"

- F approaches 0 **as $q(s)$ become more similar to $p(s|o)$**
  - If $q(s) = p(s|o), F$ orresponds only to the **surprise** under the model (low surprise = high **model evidence**)

Exact Bayesian Inference

Prior $p(s)$

$\begin{bmatrix} .5 \\ .5 \end{bmatrix}$

Likelihood $p(o|s)$

$\begin{bmatrix} .8 \\ .2 \end{bmatrix}$

$\xrightarrow{p(o|s)p(s)}$

Joint $p(o,s)$

$\begin{bmatrix} .4 \\ .1 \end{bmatrix}$

$\xrightarrow{\sum_s p(o,s)}$

Often intractable

Marginal likelihood $\mathbf{p(o)}$

$[.5]$

$\xrightarrow{p(o,s)/p(o)}$

Posterior $\mathbf{p(s|o)}$

$\begin{bmatrix} .8 \\ .2 \end{bmatrix}$

Approximate Bayesian Inference

Set initial approximate posterior $q(s) = p(s)$

$$q(s) = \begin{bmatrix} .5 \\ .5 \end{bmatrix}$$

$$F = \sum_{s \in S} q(s) \ln \frac{q(s)}{p(o,s)}$$

Initial ***F***

$$F = .5\ln\frac{.5}{.4} + .5\ln\frac{.5}{.1} = .916$$

Update 1

$q(s) = \begin{bmatrix} .6 \\ .4 \end{bmatrix}$

$$F = .6\ln\frac{.6}{.4} + .4\ln\frac{.4}{.1} = .798$$

Update 2

$q(s) = \begin{bmatrix} .7 \\ .3 \end{bmatrix}$

$$F = .7\ln\frac{.7}{.4} + .3\ln\frac{.3}{.1} = 0.721$$

Update 3

$q(s) = \begin{bmatrix} .8 \\ .2 \end{bmatrix}$

$$F = .8\ln\frac{.8}{.4} + .2\ln\frac{.2}{.1} = .693$$

Update 4

$q(s) = \begin{bmatrix} .9 \\ .1 \end{bmatrix}$

$$F = .9\ln\frac{.9}{.4} + .1\ln\frac{.1}{.1} = .730$$

Complexity minus accuracy

Complexity + **Prediction error**

Complexity **Accuracy**

$$F = \boxed{D_{KL}[q(s)||p(s)]} - \boxed{E_{q(s)}[\ln p(o|s)]}$$

"How much do beliefs change?"

"How accurate are my predictions?"

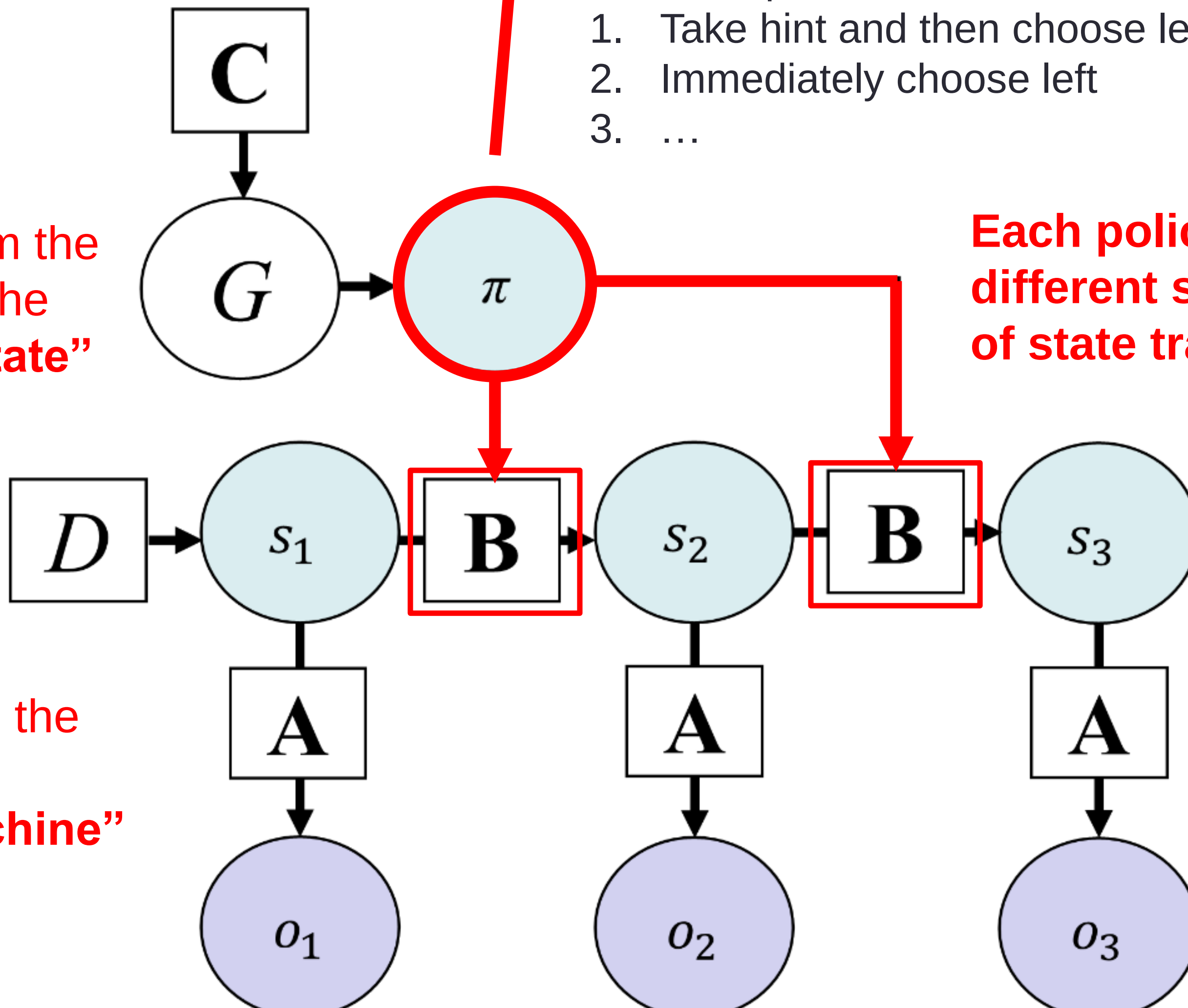
- **Complexity**: Move posterior beliefs **as little as possible from prior beliefs**
- **Accuracy**: Adjust beliefs to make the **most accurate predictions possible**

- **Convergence to $q(s)$ in message passing can be understood as prediction error minimization**

Decision-making

Policies:
possible sequences of actions

For example:
1. Take hint and then choose left machine
2. Immediately choose left
3. …



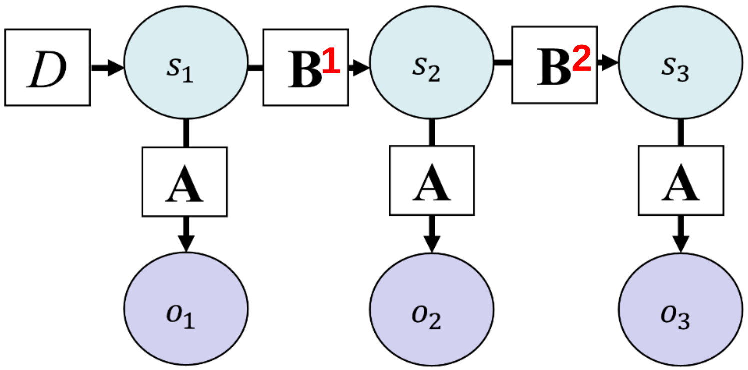
Transitioning from the **"start" state** to the **"take the hint state"**

Each policy entails a different sequence of state transitions

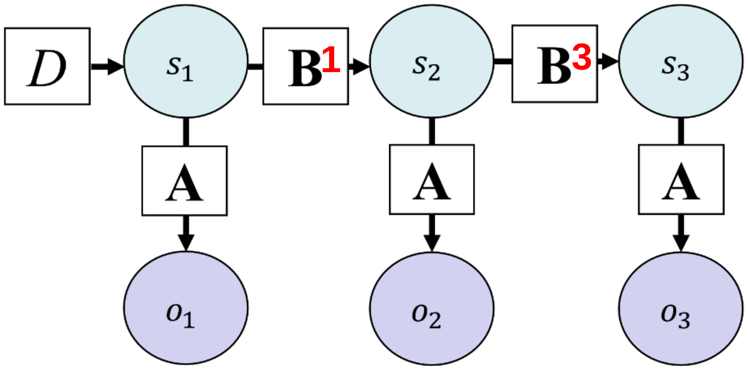
Transitioning from the "start" state to the **"choose left machine"**

Perform **message passing** for each policy in parallel

Policy 1



Policy 2



$$P(Action|\alpha) = \sigma(\alpha \times \ln P(Action|\pi))$$

α controls **randomness** in choice

Beliefs about transitions (actions), $P(s_{t+1}|s_t,\pi)$

Task instructions: "On each trial you have the choice of two slot machines. One will pay out 4 dollars, and the other will pay out \$0. You can either guess right away or you can ask for a hint about which one is more likely to pay out. However, if you ask for the hint, you can only win 2 dollars."

Take the hint

$$p(s_{\tau+1}^{choice}|s_{\tau}^{choice}, U = get\ hint) =$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Choose the left machine

$$p(s_{\tau+1}^{choice}|s_{\tau}^{choice}, U = choose\ left) =$$

$$\begin{array}{c} \begin{array}{cccc} \text{S} & \text{H} & \text{L} & \text{R} \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

Choose the right machine

$$p(s_{\tau+1}^{choice}|s_{\tau}^{choice}, U = choose\ right) =$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Example states and outcomes

Hidden State Factor 1: Context
- Left machine is more likely to pay out
- Right machine is more likely to pay out

Hidden State Factor 2: Choice states
- Start
- Ask for the hint
- Choose left machine
- Choose right machine

Outcome Modality 1: Hint
- No hint
- Left machine is more likely
- Right machine is more likely

Outcome Modality 2: Outcome
- Start
- Lose
- Win

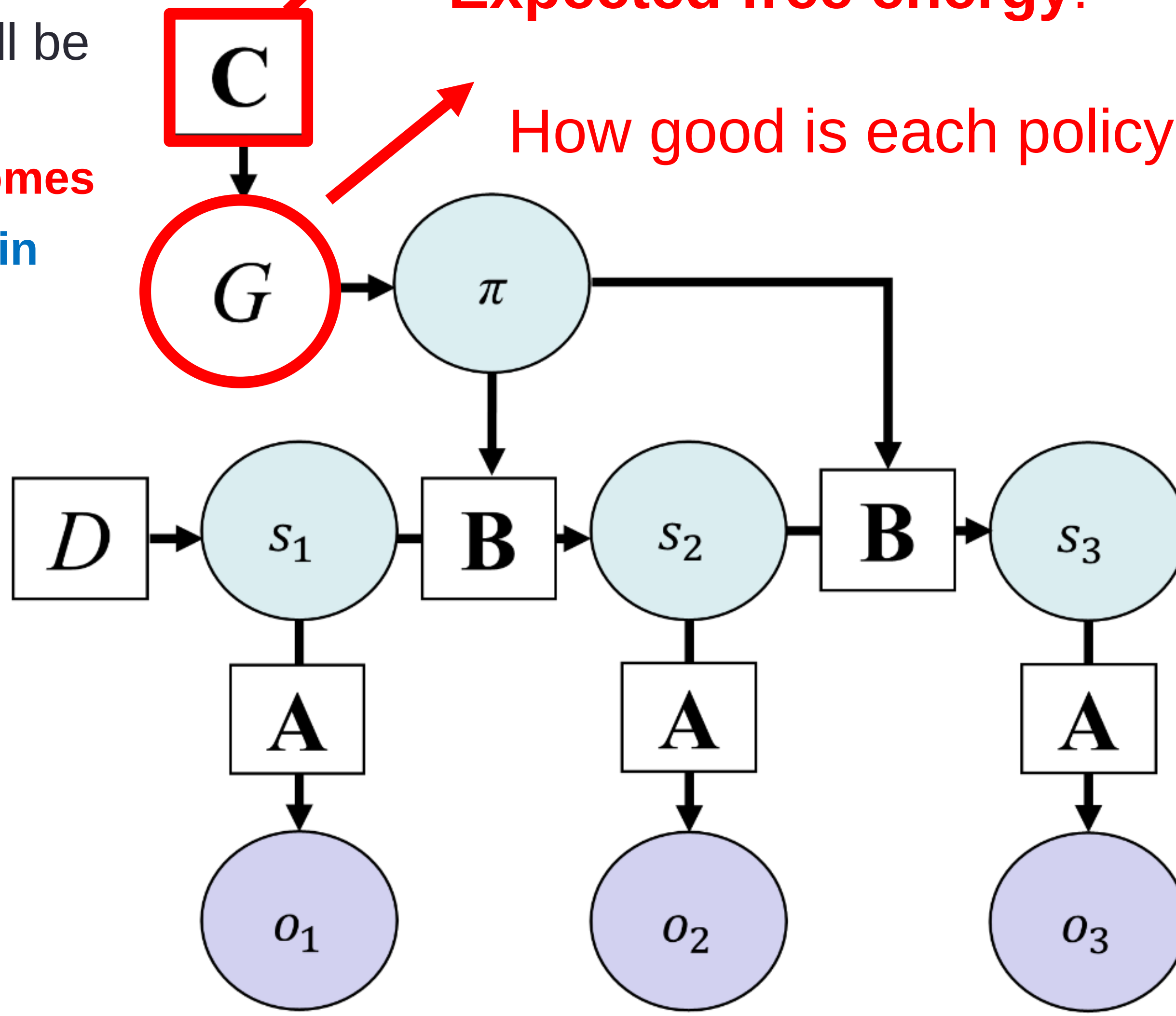
Policies with the **lowest expected free energy** will be those expected to:

- **generate preferred outcomes**
- **maximize information gain**

Preferred observations (reward)

Expected free energy:

How good is each policy?



Preference precision

$C^{win}=$

Winning is twice as preferred at T2 as at T3

$$
\mathbf{C} = \begin{array}{c} \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & 4 & 2 \end{array}\right] \begin{array}{l} \text{Start} \\ \text{Lose} \\ \text{Win} \end{array} \\ \begin{array}{ccc} \text{T1} & \text{T2} & \text{T3} \end{array} \end{array}
$$

$$
\boxed{\ln(\sigma(\mathbf{C}))} = \begin{bmatrix} -1.1 & -4.0 & -2.1 \\ -1.1 & -5.0 & -3.2 \\ -1.1 & -0.02 & -0.2 \end{bmatrix}
$$

Log probabilities

Expected free energy

- Variational free energy (F_π) pertains to **current** observations

- However, decision-making requires **predicting future observations under each policy**
  - $q(o|\pi)$, where o is treated as a random variable

- Then calculating the **expected free energy (G_π)** associated with each policy based on those expected observations

  - This is used as a means of **scoring which policies are better than others**

C = preferred/rewarding outcomes

Expected free energy

- The literature shows many different decompositions:

"Epistemic value" "Pragmatic value"

$$G_\pi = -\boxed{E_{q(o,s|\pi)}[\ln q(s|\underline{o},\pi) - \ln q(s|\pi)]} - \boxed{E_{q(o|\pi)}[\ln p(o|C)]}$$

"How much do I **expect** beliefs **will** change after a new observation **if** I choose this policy?"

Expected probability of **preferred outcomes**

"Risk" "Ambiguity"

$$= \boxed{D_{KL}[q(\underline{o|\pi})||p(\underline{o|C})]} + \boxed{E_{q(s|\pi)}[H[p(o|s)]]}$$

Expected difference between **predicted and preferred outcomes**

"How **informative** do I expect an observation will be?"

$$\begin{array}{c c c c} & \boxed{S1} & S2 & \\ p(o|s) = & \begin{bmatrix} .9 & .2 \\ .1 & .8 \end{bmatrix} & & \begin{array}{c} O1 \\ O2 \end{array} \end{array}$$

Less ambiguous

$$Entropy =$$

$$H[p(x)] = -\sum_x p(x)\ln p(x)$$

Expected free energy

- Provides a principled approach to arbitrating the **explore-exploit dilemma**:

“Risk” “Ambiguity”

$$G_\pi = \boxed{D_{KL}[q(o|\pi)||p(o|C)]} + \boxed{\mathrm{E}_{q(s|\pi)}[\mathrm{H}[p(o|s)]]}$$

Minimize deviation from preferred outcomes

Minimize uncertainty

- **Policy selection is initially information-seeking**
  - ambiguity term dominates

- Then the agent **leverages that information** to **bring about preferred outcomes**

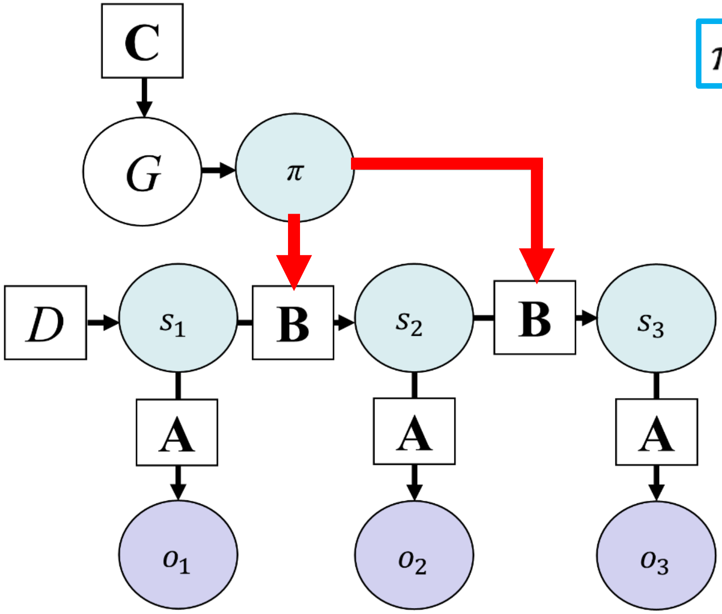
Graphical model of POMDP structure

- Each policy entails **a different set of state transitions**
- This in turn **predicts different sequences of observations**
- These jointly allow evaluation of G_π

"Risk" "Ambiguity"

$$G_\pi = \sum_\tau \boxed{\left(\mathbf{A}s_{\pi,\tau} \cdot \left(\ln \mathbf{A}s_{\pi,\tau} - \ln \mathbf{C}_\tau\right)\right.} - \boxed{\left. diag\left(\mathbf{A}^{\mathrm{T}} \ln \mathbf{A}\right) \cdot s_{\pi,\tau}\right)}$$

$$\boxed{\pi = \sigma(-G)}$$



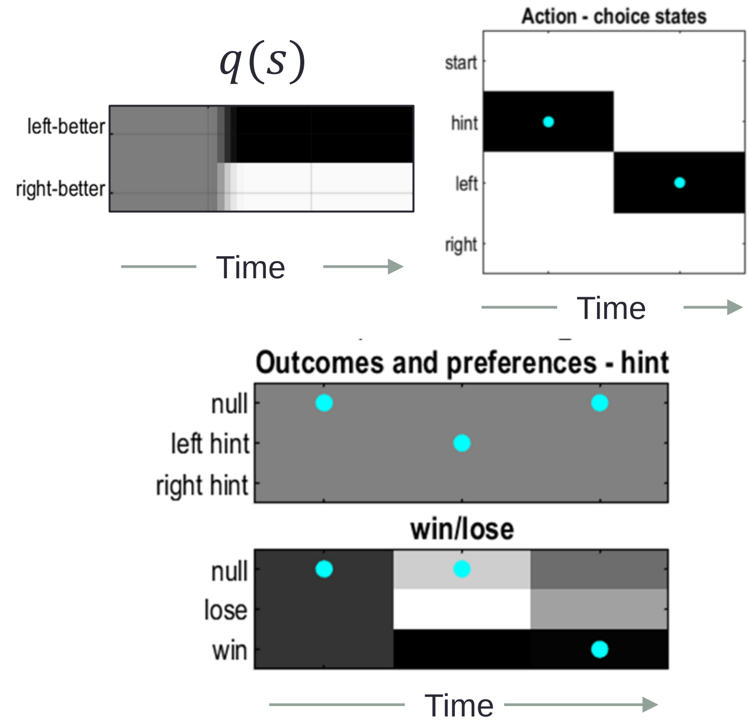
Low G_π = high probability of choosing that policy

Importantly, the **precision of $p(o|C)$** effectively weights this tradeoff and can make agents **more or less risk-seeking**

Task instructions: "On each trial you have the choice of two slot machines. One will pay out 4 dollars, and the other will pay out $0. You can either guess right away or you can ask for a hint about which one is more likely to pay out. However, if you ask for the hint, you can only win 2 dollars."

C = **moderately** precise preference distribution

Took the hint and then chose the left machine

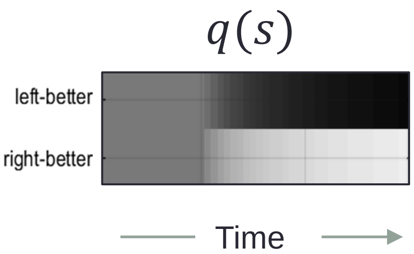


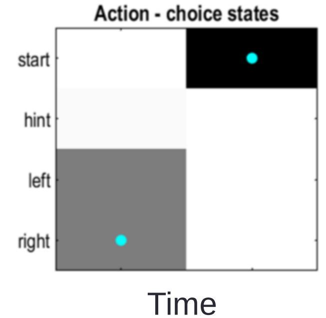
Darker = higher probability
Cyan dot = true state/outcome/action

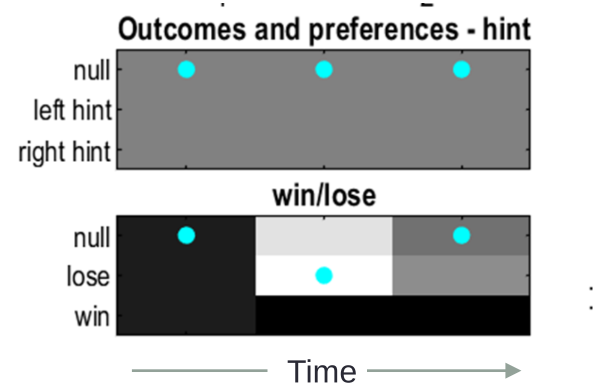
Task instructions: “On each trial you have the choice of two slot machines. One will pay out 4 dollars, and the other will pay out $0. You can either guess right away or you can ask for a hint about which one is more likely to pay out. However, if you ask for the hint, you can only win 2 dollars.”

C = **very** precise preference distribution

Took a guess immediately and lost







Darker = higher probability
Cyan dot = true state/outcome/action

Increased exploitation with learning

- Each trial where it **infers the left machine is better**, it **increases its prior belief** that the left machine will be better in the future

- This elevated confidence **reduces the drive to seek information**

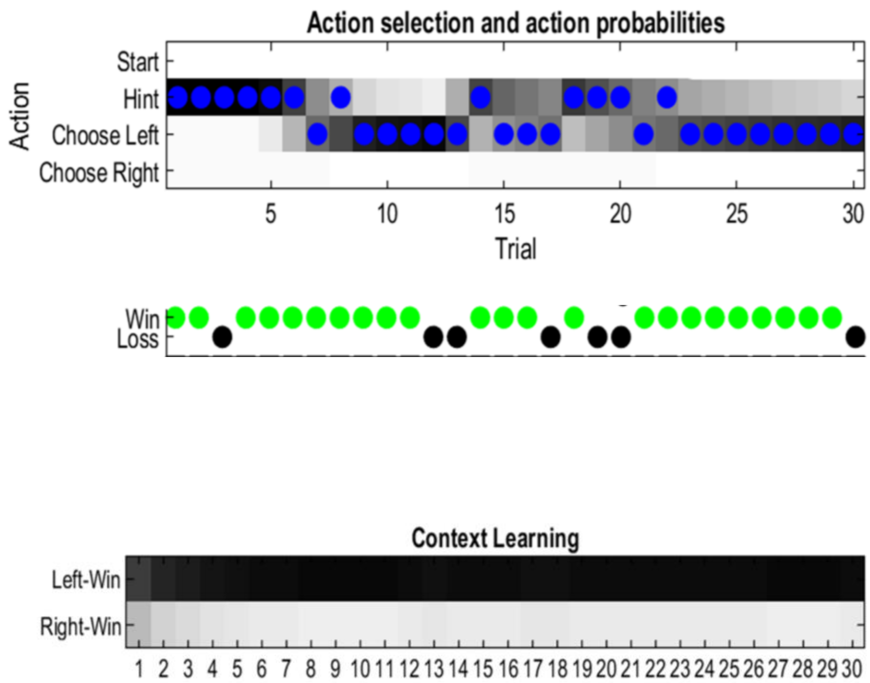
- Preference precision also **moderates the influence of action on these dynamics**

- Consider both **stable** and **volatile** environments

Increased exploitation with learning

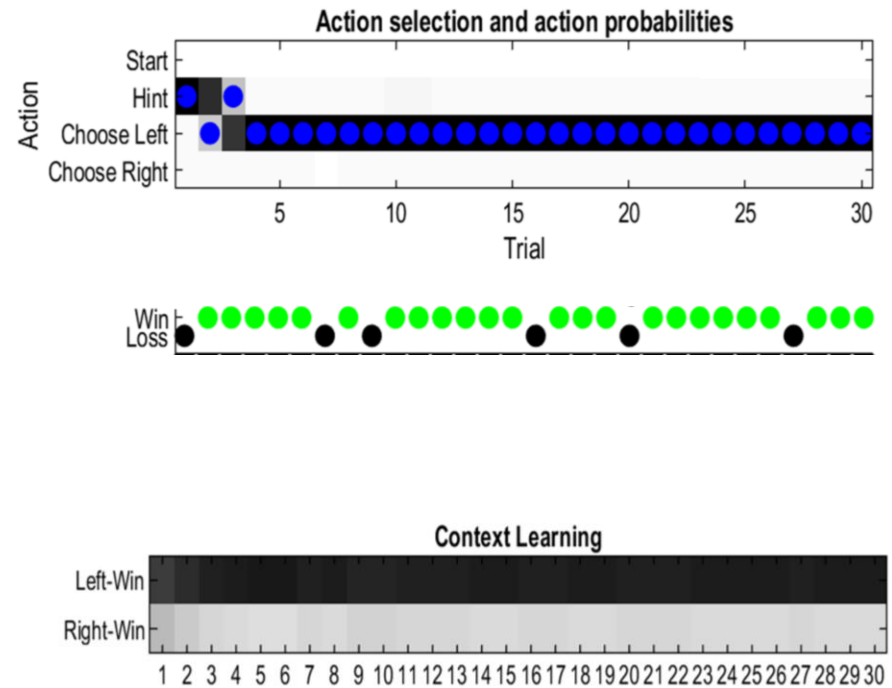
C = **moderately** precise preference distribution

Risk-Averse Agent



C = **very** precise preference distribution

Risk-Seeking Agent



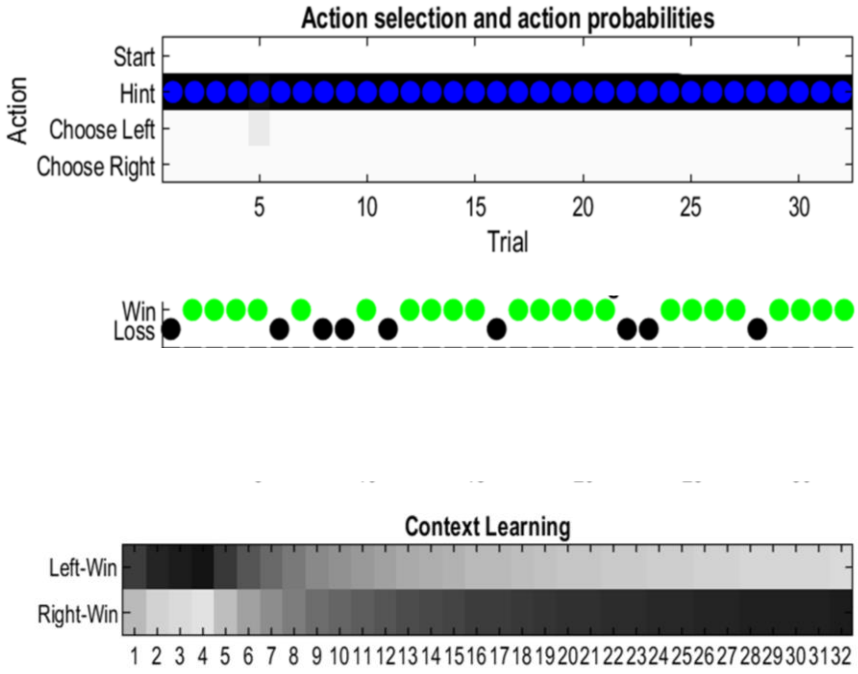
Left machine is always better

Darker = higher probability
Blue circle = first action
Green = win
Black = loss

Increased exploitation with learning

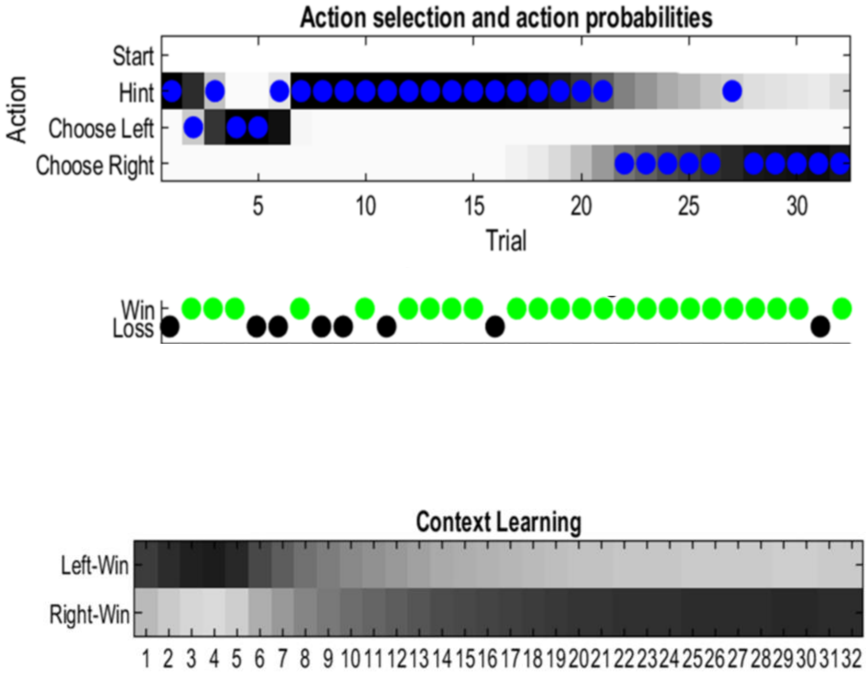
C = **moderately** precise preference distribution

Risk-Averse Agent



C = **very** precise preference distribution

Risk-Seeking Agent

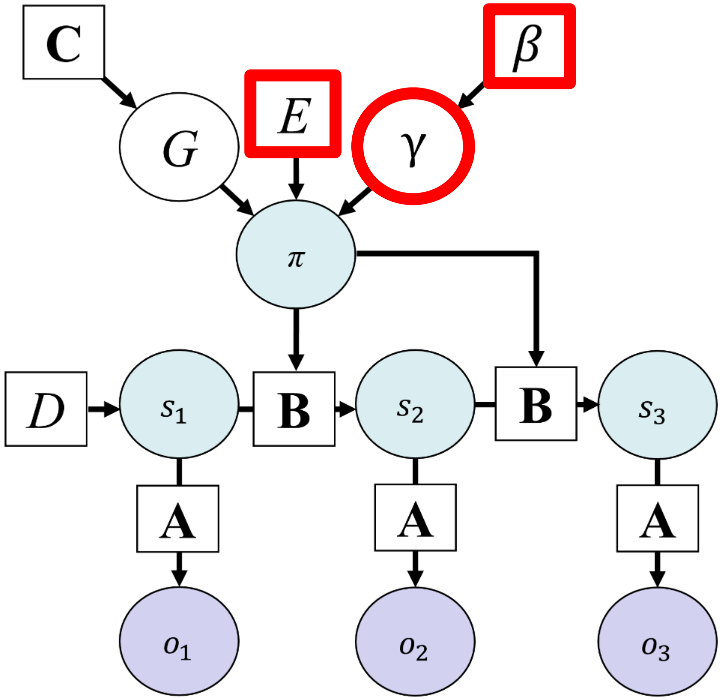


Right machine unexpectedly becomes better

Darker = higher probability
Blue circle = first action
Green = win
Black = loss

Graphical model of POMDP structure

- The vector E is prior over policies $p(\pi)$, which allows one to model **habits** and **habit learning**
- γ is the expected precision of G
    - An **inverse temperature parameter** based on **confidence** in the agent's action model
    - Lower confidence **increases the influence of habits**



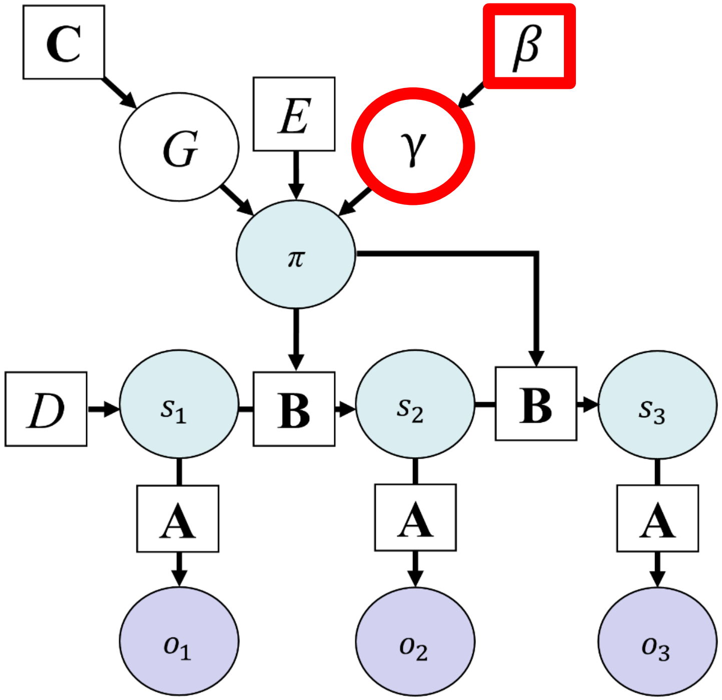
$$\pi = \sigma(\ln E - \gamma G)$$

$$p(\gamma) = \Gamma(1, \beta)$$

$$E[\gamma] = \gamma = 1/\beta$$

Graphical model of POMDP structure

- γ can be **updated** based on changes in confidence **after new observations**
  - These updates have been proposed to capture **phasic dopamine responses** within the **neural process theory**



$$\pi_0 = \sigma(\ln E - \gamma G)$$

$$\pi = \sigma(\ln E - \boxed{F} - \gamma G)$$

$$\beta_{update} = \beta - \beta_0 + \boxed{(\pi - \pi_0) \cdot (-G)}$$

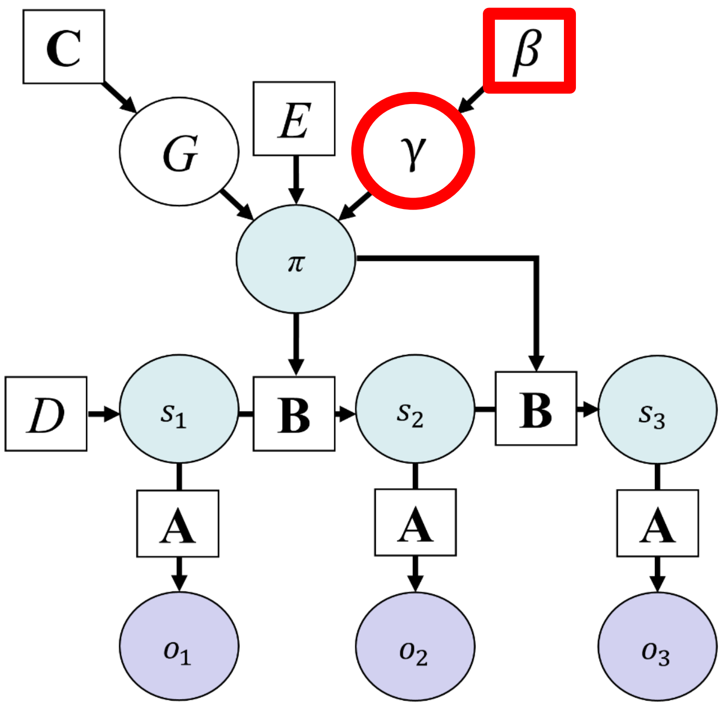
$$\beta = \beta - \beta_{update}$$

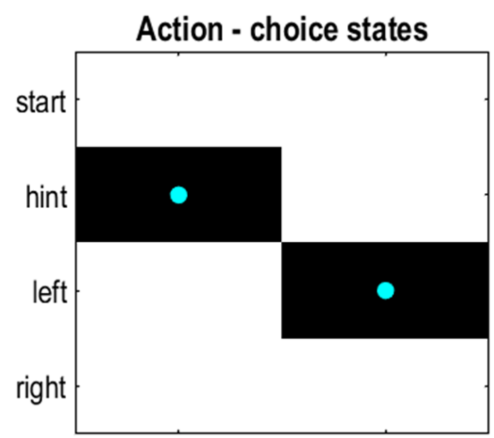
$$\gamma = 1/\beta$$

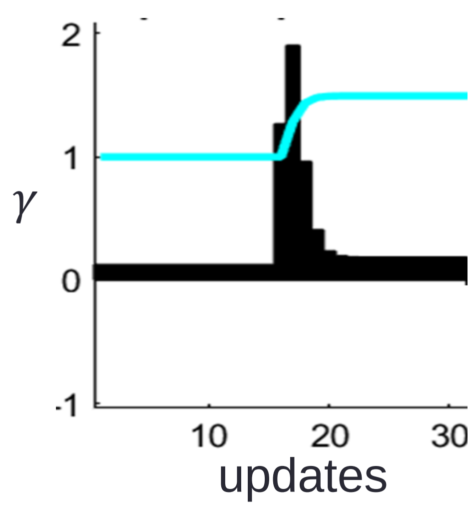
Were new observations **consistent with prior beliefs about G**?

Graphical model of POMDP structure

- Example of γ updating in the **slot machine task**
  - Agent **expected** that **taking the hint was better**
  - After taking the hint, it became **more confident in that policy**
    - This is **evidence that G was reliable**, and so γ **increases**







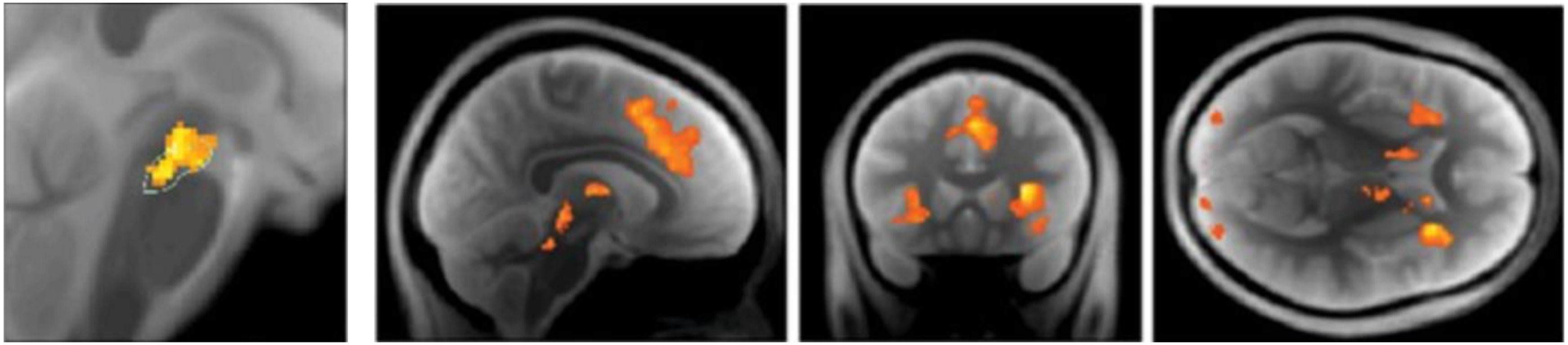
Black = rate of change in γ

Neural process theory

- One study has shown **that γ updating is correlated with neural activity in midbrain dopamine regions** as well as dopaminergic targets
  - Theory remains to be thoroughly tested

The Dopaminergic Midbrain Encodes the Expected Certainty about Desired Outcomes

Philipp Schwartenbeck, Thomas H. B. FitzGerald, Christoph Mathys, Ray Dolan, Karl Friston Author Notes



Active Learning

- Two types of exploration:
  - **State exploration:** seeking information about hidden states
    - Left machine better vs. right machine better (take hint)

  - **Parameter exploration**: seeking information about parameters
    - **Choosing the left machine** to **resolve uncertainty about the probability of a win,** $\boldsymbol{p(o|s)}$
      - In other words, **choose the machine for which you are least confident about the outcomes it will produce**

  - Driven by an **added term to the expected free energy** when learning is included
    - Based on using **Dirichlet distributions** as priors on the categorical distributions described so far.

Learning

$$p(\boldsymbol{\theta}|\boldsymbol{\alpha}) = Dir(\boldsymbol{\theta}|\boldsymbol{\alpha}) = \frac{\Gamma(\sum_{k=1}^{K}\alpha_k)}{\prod_{k=1}^{K}\Gamma(\alpha_k)}\prod_{k=1}^{K}\theta_k^{\alpha_k-1}$$

- Dirichlet distributions are **conjugate priors** for categorical distributions
  - This means that inference using a **Dirichlet distribution as a prior**
  - and a **categorical distribution as a likelihood**
  - will return a **Dirichlet distribution as a posterior**
    - In other words, you can **keep learning** and the **form of the distribution** doesn't change

$$p(\mathbf{A}) = Dir(\boldsymbol{a}) \longleftarrow$$

The probability that **A** has a particular shape

$$\boldsymbol{a} = p(o_\tau|\, s_\tau) = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \\ a_5 & a_6 \end{bmatrix}$$

Concentration parameters ("counts")

Learning

$$p\left(o_\tau^{win} \mid s_\tau^{context,choice=choose\ left}\right) =$$

$$\boldsymbol{a} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{.2} & \mathbf{.8} \\ \mathbf{.8} & \mathbf{.2} \end{bmatrix}$$

$$p\left(o_\tau^{win} \mid s_\tau^{context,choice=choose\ left}\right) = \boldsymbol{a} = \begin{array}{c c} \text{Left} & \text{Right} \\ \left[\begin{matrix} a_1 & a_2 \\ a_3 & a_4 \\ a_5 & a_6 \end{matrix}\right. & \left.\begin{matrix} \\ \\ \\ \end{matrix}\right] \end{array} \begin{array}{l} \\ \text{Start} \\ \text{Lose} \\ \text{Win} \end{array}$$

$$\boldsymbol{a}_{trial+1} = \boldsymbol{a}_{trial} + \sum_\tau o_\tau \otimes s_\tau$$

$$p\left(o_\tau^{win} \mid s_\tau^{context,choice=choose\ left}\right) =$$

$$p(s_\tau^{context}) = [0 \quad 1]$$

$$\boldsymbol{a} = \begin{array}{cc} \text{Left} & \text{Right} \\ \end{array} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0.25} & \mathbf{0.25} \\ \mathbf{0.25} & \mathbf{0.25} \end{bmatrix} \begin{array}{l} \text{Start} \\ \text{Lose} \\ \text{Win} \end{array} \longrightarrow \underbrace{\begin{bmatrix} \mathbf{0} \\ \mathbf{1} \\ \mathbf{0} \end{bmatrix}}_{o_\tau^{win}} \longrightarrow \boldsymbol{a} = \begin{array}{cc} \text{Left} & \text{Right} \\ \end{array} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0.25} & \mathbf{1.25} \\ \mathbf{0.25} & \mathbf{0.25} \end{bmatrix} \begin{array}{l} \text{Start} \\ \text{Lose} \\ \text{Win} \end{array}$$

Concentration parameters

Learning

$$p(o_\tau^{win} | s_\tau^{context,choice=choose\ left}) =$$

$$\boldsymbol{a} = \begin{bmatrix} 0 & 0 \\ .2 & .8 \\ .8 & .2 \end{bmatrix}$$

$$p(o_\tau^{win} | s_\tau^{context,choice=choose\ left}) = \boldsymbol{a} = \begin{array}{c} \\ \\ \\ \end{array}\begin{array}{cc} \text{Left} & \text{Right} \\ \end{array}$$

$$\begin{array}{cc|l} \text{Left} & \text{Right} & \\ a_1 & a_2 & \text{Start} \\ a_3 & a_4 & \text{Lose} \\ a_5 & a_6 & \text{Win} \end{array}$$

$$\boldsymbol{a}_{trial+1} = \omega \times \boldsymbol{a}_{trial} + \eta \times \sum_\tau o_\tau \otimes s_\tau$$

Forgetting rate

Learning rate

Active Learning

- Information-seeking about parameters. Expanded expected free energy
  - **Novelty** term reflecting **how much beliefs about A are expected to change** with a new observation

Risk + Ambiguity "Novelty"

$$G_\pi = D_{KL}[q(o_\tau|\pi)||p(o_\tau)] + \mathrm{E}_{q(s_\tau|\pi)}[\mathrm{H}[p(o_\tau|s_\tau)] - \mathrm{E}_{p(o_\tau|s_\tau)q(s_\tau|\pi)}[D_{KL}[q(\mathbf{A}|o_\tau, s_\tau)||q(\mathbf{A})]]$$

$$\approx \sum_\tau \left(\mathbf{A}s_{\pi,\tau} \cdot \left(\ln \mathbf{A}s_{\pi,\tau} - \ln \mathbf{C}_\tau\right) - diag\left(\mathbf{A}^\mathrm{T} \ln \mathbf{A}\right) \cdot s_{\pi,\tau} - \boxed{\mathbf{A}s_{\pi,\tau} \cdot \mathbf{W}s_{\pi,\tau}}\right)$$

$$\mathbf{W} := \frac{1}{2}\left(\boldsymbol{a}^{\odot(-1)} - \boxed{\boldsymbol{a}_{sums}^{\odot(-1)}}\right)$$

Sums of column values

Active Learning

- Intuition

Risk + Ambiguity "Novelty"

$$G_\pi = D_{KL}[q(o_\tau|\pi)||p(o_\tau)] + \mathrm{E}_{q(s_\tau|\pi)}[\mathrm{H}[p(o_\tau|s_\tau)] - \mathrm{E}_{p(o_\tau|s_\tau)q(s_\tau|\pi)}[D_{KL}[q(\mathbf{A}|o_\tau, s_\tau)||q(\mathbf{A})]]$$

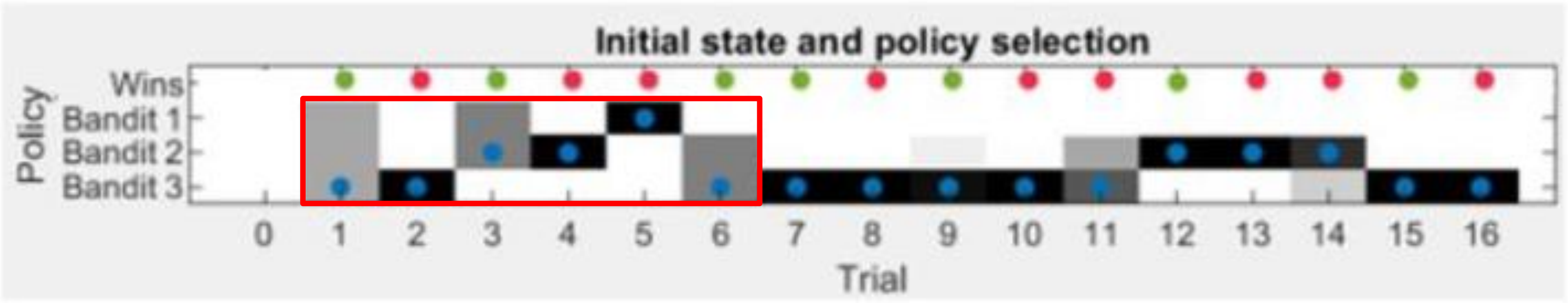
Drive toward the "Right" state

$$\boldsymbol{a} = \begin{array}{cc} \text{Left} & \text{Right} \\ \left[\begin{array}{cc} \mathbf{0} & \mathbf{0} \\ \mathbf{25.0} & \mathbf{.25} \\ \mathbf{25.0} & \mathbf{.25} \end{array}\right] \end{array} \begin{array}{l} \text{Start} \\ \text{Lose} \\ \text{Win} \end{array}$$

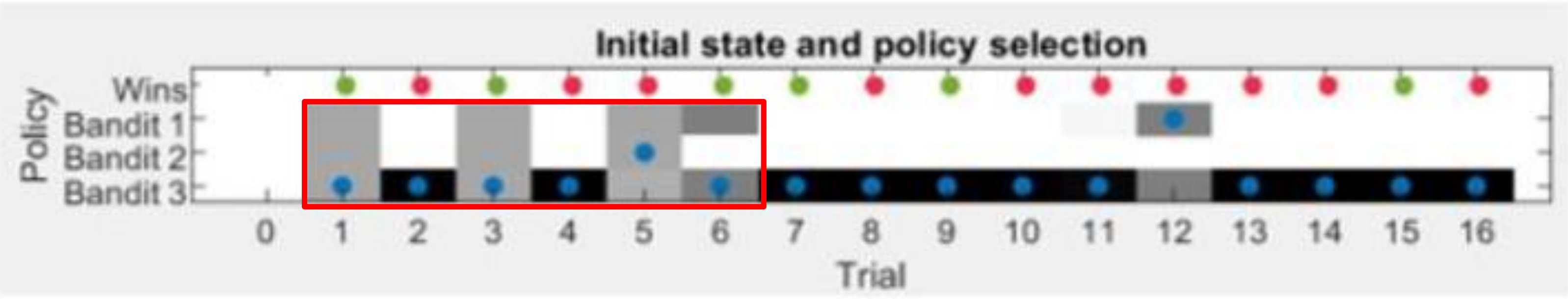
Active Learning

- Three slot machines with unknown reward probabilities

Information value included



Reward only model



Summary

- Active inference describes **perception, learning, and decision-making** in terms of approximate Bayesian inference
- It uses a **variational message passing** scheme that is **biologically plausible** and can generate neuroscientific predictions
- It includes **reward-seeking** as well as **information-seeking** drives
  - **Active learning**: Additional information-seeking drive when trying to learn model parameters (e.g., reward probabilities)

Thanks!

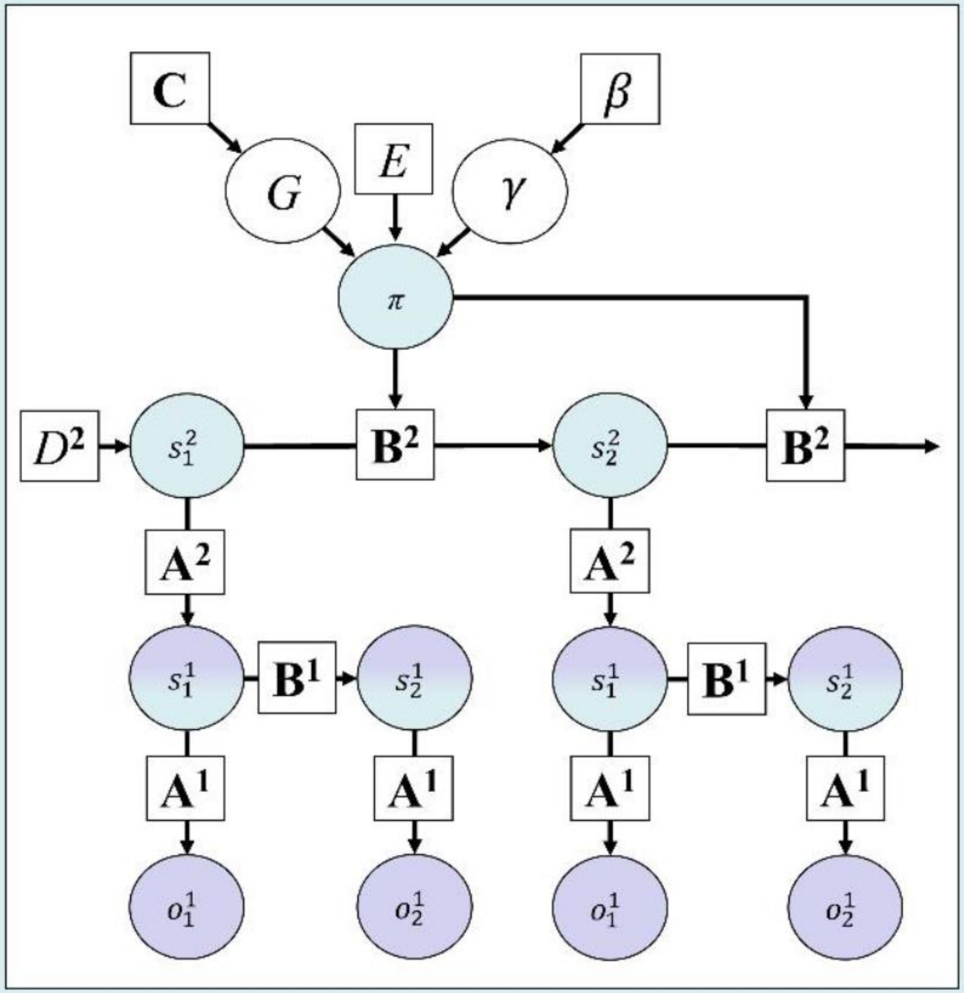
1. William K. Warren Foundation
2. National Institute for General Medical Sciences (P20GM121312)
3. National Institute of Mental Health (MH123691)

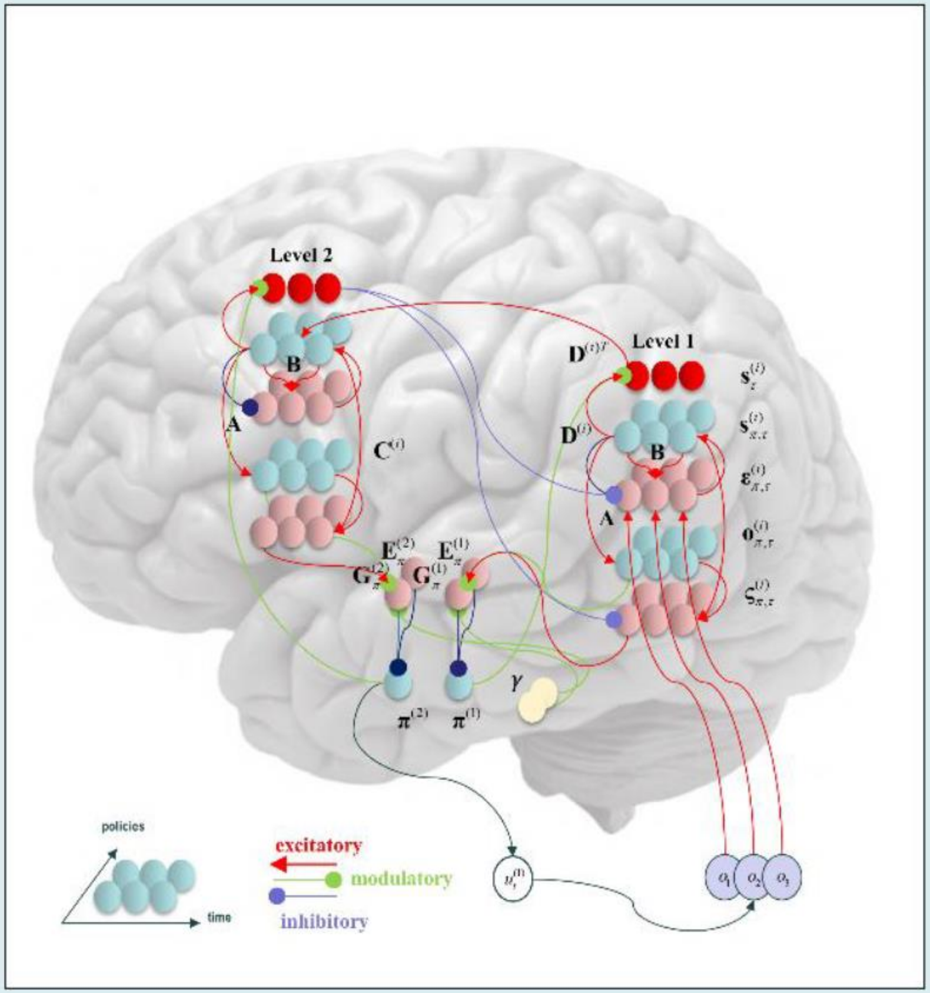
Karl Friston

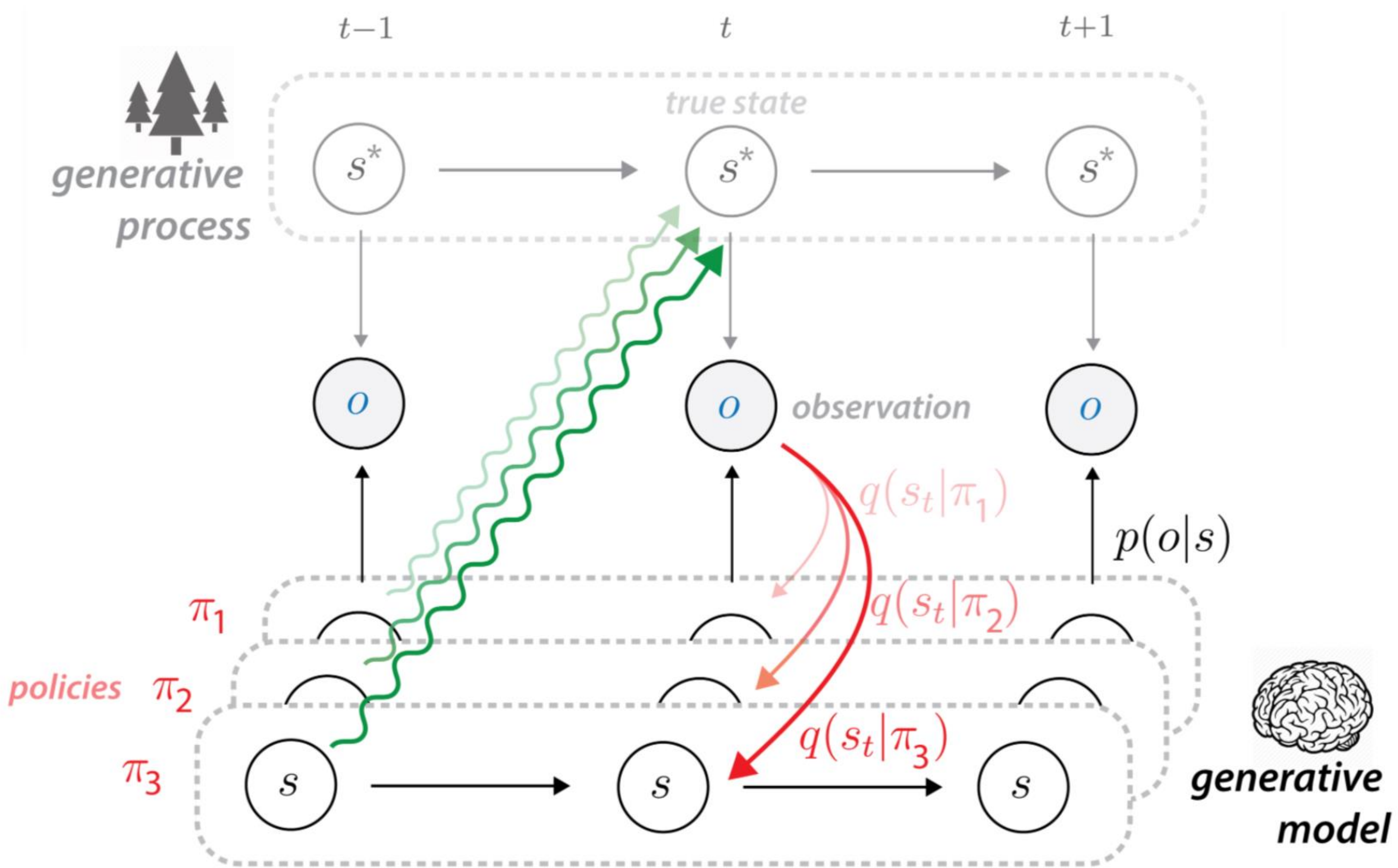
Christopher Whyte

Supplemental

Hierarchical models







Oleg Solopchuk: https://medium.com/@solopchuk/tutorial-on-active-inference-30edcf50f5dc

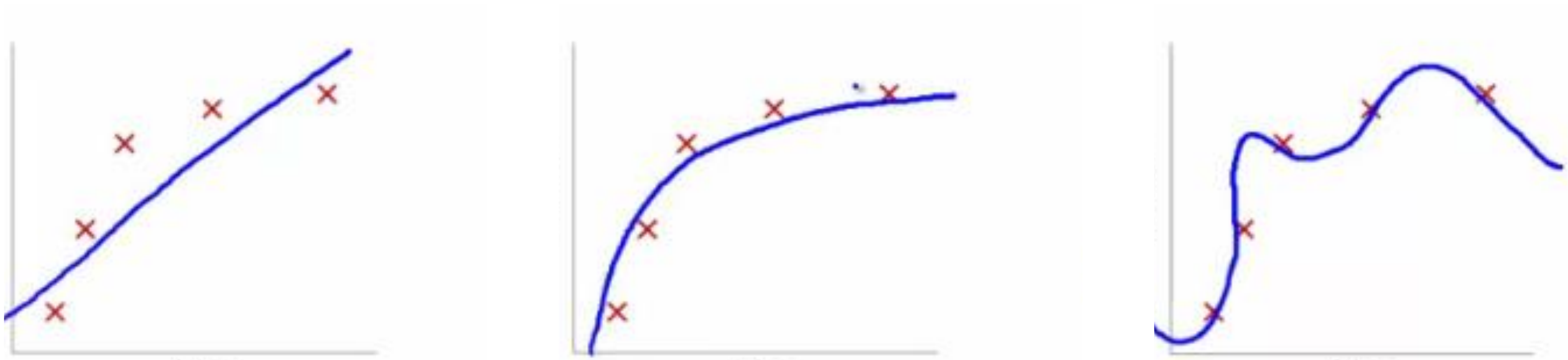
Brief clarification on policies

- In reinforcement learning, the term “policy” is often used to refer to a **mapping from states to actions**
  - For example:
    - “If I’m in state #1, then I will go left”
    - “If I’m in state #2, then I will go right”

- In active inference, a “policy” refers to possible **sequence of actions** of some pre-set depth.
  - For example:
    - Policy #1: “I will go left, then left, then right”
    - Policy #2: “I will go right, then right, then left”

    - The **value of the whole sequence** is evaluated based on the predicted sequence of observations

Why not just maximize accuracy?



$\theta_0 + \theta_1 x$

$\theta_0 + \theta_1 x + \theta_2 x^2$

$\theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$

Overfitting may minimize prediction error for existing observations, but it will not be good at making new predictions

Expected free energy

$$
\begin{aligned}
G_\pi &= \mathrm{E}_{q(o,s|\pi)}[\ln q(s|\pi) - \ln p(o,s|\pi)] \\
&= \mathrm{E}_{q(o,s|\pi)}[\ln q(s|\pi) - \ln p(s|o,\pi)] - \mathrm{E}_{q(o|\pi)}[\ln p(o|\pi)] \\
&\approx \mathrm{E}_{q(o,s|\pi)}[\ln q(s|\pi) - \ln q(s|o,\pi)] - \mathrm{E}_{q(o|\pi)}[\ln p(o|C)] \\
&= -\mathrm{E}_{q(o,s|\pi)}[\ln q(s|o,\pi) - \ln q(\mathrm{s}|\pi)] - \mathrm{E}_{q(o|\pi)}[\ln p(o|C)] \\
&= D_{KL}[q(o|\pi)||p(o|C)] + \mathrm{E}_{q(s|\pi)}[\mathrm{H}[p(o|s)]]
\end{aligned}
$$

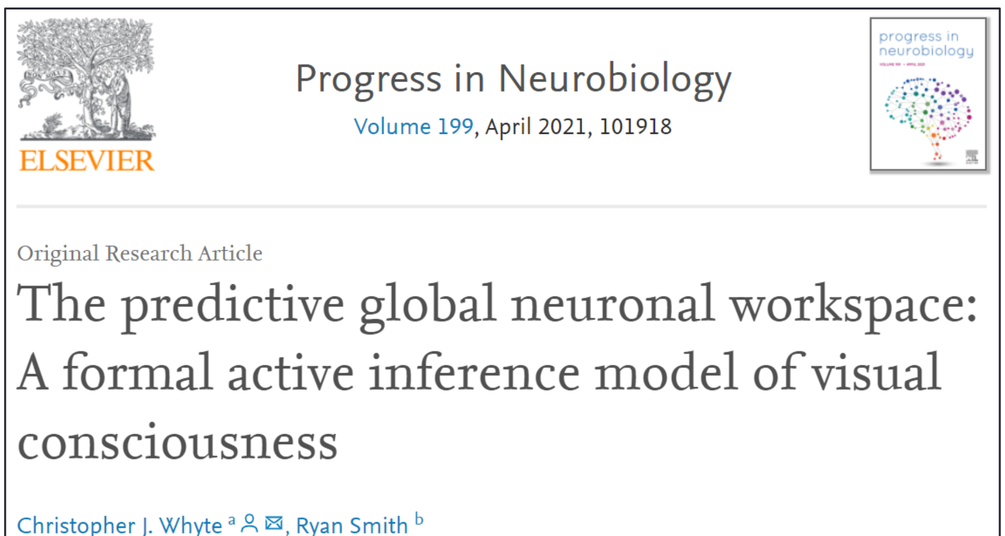
Neural process theory

- Plausible mapping from algorithm to neural implementation
- Based on **variational message passing** and the idea of treating **free energy gradients as prediction errors**
- Allows predictions that can be used in fMRI and EEG studies

Friston, K., T. FitzGerald, F. Rigoli, P. Schwartenbeck, and G. Pezzulo. **Active Inference: A Process Theory**. *Neural Computation* 29, no. 1 (2017): 1-49.

Neural process theory

- The **rate of change in posterior beliefs (*s*)** corresponds to empirically observed **ERPs** in several tasks. For example:
  - Inattentional blindness tasks



Progress in Neurobiology

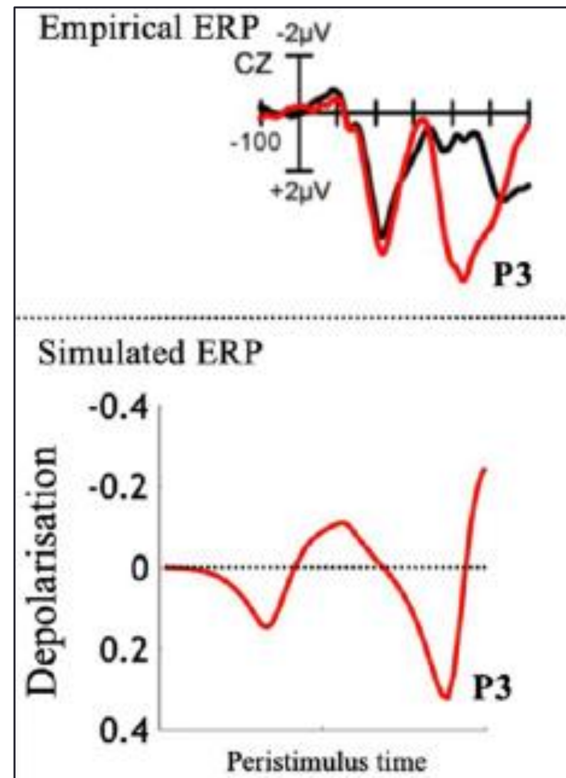
Volume 199, April 2021, 101918

ELSEVIER

Original Research Article

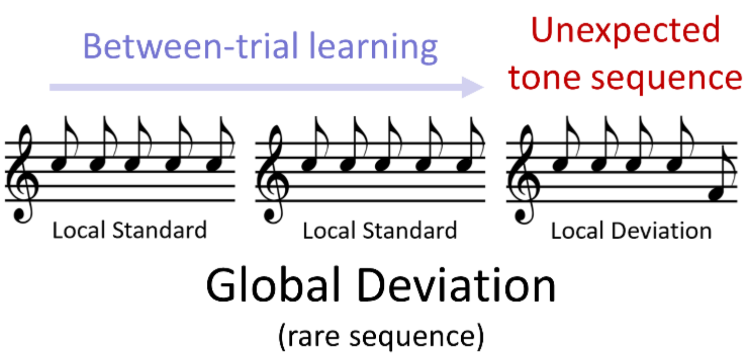
The predictive global neuronal workspace: A formal active inference model of visual consciousness

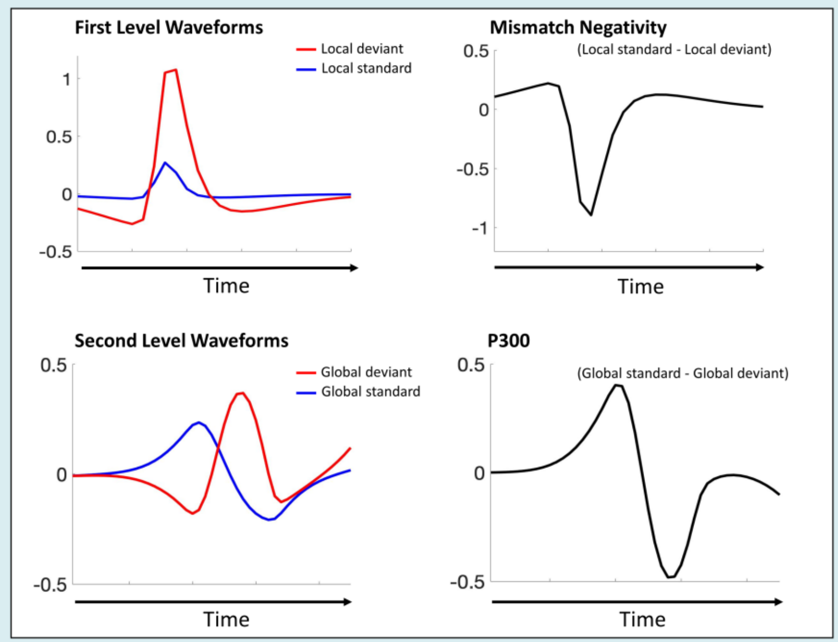
Christopher J. Whyte [a], Ryan Smith [b]



Neural process theory

- The **rate of change in posterior beliefs (*s*)** corresponds to empirically observed **ERPs** in several tasks. For example:
  - Inattentional blindness tasks
  - Oddball tasks





(From tutorial paper)

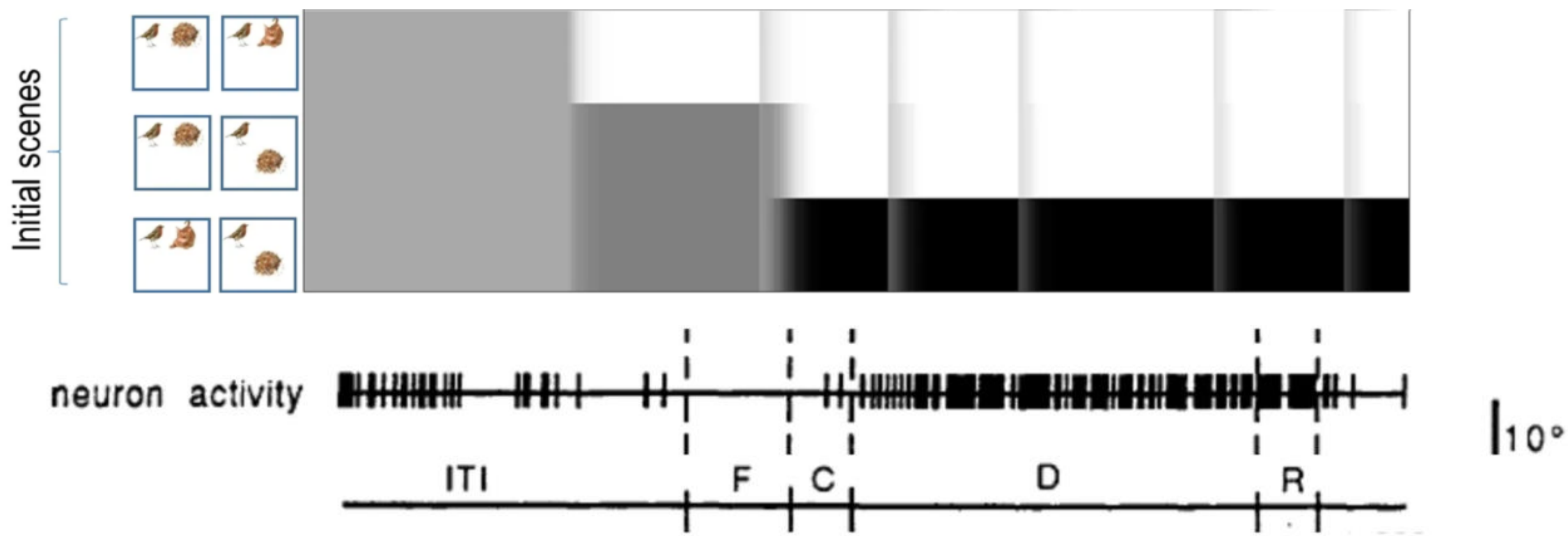
Neural process theory

- **Posterior beliefs (*s*)** corresponds to empirically observed **firing rates** in ‘delay neurons’ within working memory tasks

Working memory, attention, and salience in active inference

Thomas Parr & Karl J Friston

Scientific Reports **7**, Article number: 14678 (2017) | Cite this article



Neural process theory

$$\varepsilon_{\pi,\tau} \leftarrow \frac{1}{2}\left(\ln\left(\mathbf{B}_{\pi,\tau-1} s_{\pi,\tau-1}\right) + \ln\left(\mathbf{B}^{\dagger}_{\pi,\tau+1} s_{\pi,\tau+1}\right)\right) + \ln \mathbf{A}^{\mathrm{T}} o_\tau - \ln s_{\pi,\tau}$$

$$v_{\pi,\tau} \leftarrow \ln s_{\pi,\tau}; \longrightarrow v_{\pi,\tau} \leftarrow v_{\pi,\tau} + \varepsilon_{\pi,\tau}; \longrightarrow s_{\pi,\tau} \leftarrow \sigma(v_{\pi,\tau})$$

Iterated to convergence

$$s_\tau = \sum\nolimits_\pi \pi \cdot s_{\pi,\tau}$$

$$\mathbf{A} s_{\pi,\tau} = o_{\pi,\tau}$$

$$\varsigma_{\pi,\tau} = \mathbf{A} s_{\pi,\tau} \cdot (\ln \mathbf{A} s_{\pi,\tau} - \ln \mathbf{C}_\tau) - diag\left(\mathbf{A}^{\mathrm{T}} \ln \mathbf{A}\right) \cdot s_{\pi,\tau} - \mathbf{A} s_{\pi,\tau} \cdot \mathbf{W} s_{\pi,\tau}$$

$$\mathbf{W} := \frac{1}{2}\left(\boldsymbol{a}^{\odot(-1)} - \boldsymbol{a}^{\odot(-1)}_{sums}\right)$$

$$F_\pi = \sum_\tau s_{\pi,\tau} \cdot \varepsilon_{\pi,\tau}; \quad G_\pi = \sum_\tau \varsigma_{\pi,\tau}$$

Belief updating
(Neuronal activation dynamics)

Policy inference and expected free energy precision (Dopamine)

$$\pi_0 \leftarrow \sigma(\ln E - \gamma G); \quad \pi \leftarrow \sigma(\ln E - F - \gamma G); \longleftarrow \gamma \leftarrow 1/\beta$$

$$G_{error} \leftarrow (\pi - \pi_0) \cdot (-G); \longrightarrow \beta_{update} \leftarrow \beta - \beta_0 + G_{error} \longrightarrow \beta \leftarrow \beta - \beta_{update}$$

Iterated to convergence

$$p(u_t) = \sigma(\alpha \times \ln p(u_t|\pi))$$

Action selection

$$d_{trial+1} = \omega \times d_{trial} + \eta \times s_{\tau=1}$$

$$\boldsymbol{a}_{trial+1} = \omega \times \boldsymbol{a}_{trial} + \eta \times \sum\nolimits_\tau o_\tau \otimes s_\tau$$

Learning
(Synaptic Plasticity)